

QY	421	GTGGTGATCCGGTCCGGCGCGTGTGCAGGATTCACGCTGTTTCCGCGCTGTGGCGCG	480
Db	70685	GTGGTGATCCGGTCCGGCGCGTGTGCAGGATTCACGCTGTTTCCGCGCTGTGGCGCG	70744
QY	481	ATGCCAATGGTGTGTGCGCCGGAAAGCTATCTCACACATACGGCATACCGGAAAA	540
Db	70745	ATGCCAATGGTGTGTGCGCCGGAAAGCTATCTCACACATACGGCATACCGGAAAA	70804
QY	541	CCGCCGATTTGAGTAGTCAATTCATGGCTTGAATACAGCGTCGGCCCGACAATGAATTT	600
Db	70805	CCGCCGATTTGAGTAGTCAATTCATGGCTTGAATACAGCGTCGGCCCGACAATGAATTT	70864
QY	601	GAATGATCCACCGGAAGGGATCTCGACTCGCTGATCCCAAGGAAGATTGTGACT	660
Db	70865	GAATGATCCACCGGAAGGGATCTCGACTCGCTGATCCCAAGGAAGATTGTGACT	70924
QY	661	AATGATCCGATGACGCTGTGTGGCTGTGCGTACGGGGGTGCCGGATCGCCTACGTGCCG	720
Db	70925	AATGATCCGATGACGCTGTGTGGCTGTGCGTACGGGGGTGCCGGATCGCCTACGTGCCG	70984
QY	721	CTGATGTGGTGATCAACAGAGATCAATCGTGGGGAGCTGGAGATCCTGCTGCCGCTTAC	780
Db	70985	CTGATGTGGTGATCAACAGAGATCAATCGTGGGGAGCTGGAGATCCTGCTGCCGCTTAC	71044
QY	781	CAGTCAGATCCAGCCCGGTTTTATGGGTTATATACGAAAAAGATAAGTGGCGCTGAAG	840
Db	71045	CAGTCAGATCCAGCCCGGTTTTATGGGTTATATACGAAAAAGATAAGTGGCGCTGAAG	71104
QY	841	GTACAGGTCGTGATCAACTCGCTGACGGATTATTTGTTGAGGTCGGTAAATGTTTCAG	900
Db	71105	GTACAGGTCGTGATCAACTCGCTGACGGATTATTTGTTGAGGTCGGTAAATGTTTCAG	71164
QY	901	GAGATCCACGGGCGGGGAAAGAGAGTAA	930
Db	71165	GAGATCCACGGGCGGGGAAAGAGAGTAA	71194

RESULT 2
U00096 33
WPCOMMENT
Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End	Accession
U00096_00	1	110000	U00096_00
U00096_01	100001	210000	U00096_01
U00096_02	200001	310000	U00096_02
U00096_03	300001	410000	U00096_03
U00096_04	400001	510000	U00096_04
U00096_05	500001	610000	U00096_05
U00096_06	600001	710000	U00096_06
U00096_07	700001	810000	U00096_07
U00096_08	800001	910000	U00096_08
U00096_09	900001	1010000	U00096_09
U00096_10	1000001	1110000	U00096_10
U00096_11	1100001	1210000	U00096_11
U00096_12	1200001	1310000	U00096_12
U00096_13	1300001	1410000	U00096_13
U00096_14	1400001	1510000	U00096_14
U00096_15	1500001	1610000	U00096_15
U00096_16	1600001	1710000	U00096_16
U00096_17	1700001	1810000	U00096_17
U00096_18	1800001	1910000	U00096_18
U00096_19	1900001	2010000	U00096_19
U00096_20	2000001	2110000	U00096_20
U00096_21	2100001	2210000	U00096_21
U00096_22	2200001	2310000	U00096_22
U00096_23	2300001	2410000	U00096_23
U00096_24	2400001	2510000	U00096_24
U00096_25	2500001	2610000	U00096_25
U00096_26	2600001	2710000	U00096_26
U00096_27	2700001	2810000	U00096_27
U00096_28	2800001	2910000	U00096_28
U00096_29	2900001	3010000	U00096_29

```

Db      88262 CTGATCGGTGATCAACAGAGATCAATCGTGGGAGCTGGAGATCCTGTCGCCGTTTAC 88321
QY      781  CAGTCAGATCCAGCCGCGTTTATGCGTTATATACCGAAAGATAAGTCGCGTGAAG 840
Db      88322 CAGTCAGATCCAGCCGCGTTTATGCGTTATATACCGAAAGATAAGTCGCGTGAAG 88381
QY      841  GTACAGGTGCGTATCAATCGCTGACGGATTTATTTGTGAGTTCGGTAATGTTTCAG 900
Db      88382 GTACAGGTGCGTATCAATCGCTGACGGATTTATTTGTGAGTTCGGTAATGTTTCAG 88441
QY      901  GAGATGCACGGCGCGGGAAGAGAGTAA 930
Db      88442 GAGATGCACGGCGCGGGAAGAGAGTAA 88471

RESULT 3
AP002564
LOCUS   AP002564 307962 bp DNA linear BCT 20-MAR-2004
DEFINITION Escherichia coli O157:H7 DNA, complete genome, section 15/20.
ACCESSION AP002564 BA000007
VERSION   AP002564.1 GI:13363382
KEYWORDS
SOURCE   Escherichia coli O157:H7
ORGANISM Escherichia coli O157:H7
REFERENCE
AUTHORS  Makino,K., Yokoyama,K., Kubota,Y., Yutsudo,C.H., Kimura,S.,
Kurokawa,K., Ishii,K., Hattori,M., Tatsuno,I., Abe,H., Iida,T.,
Yamamoto,K., Ohnishi,M., Hayashi,T., Yasunaga,T., Honda,T.,
Sasakawa,C. and Shinagawa,H.
TITLE    Complete nucleotide sequence of the prophage VT2-Sakai carrying the
verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
derived from the Sakai outbreak
JOURNAL  Genes Genet. Syst. 74 (5), 227-239 (1999)
MEDLINE  20198780
PUBMED   10734605
REFERENCE
AUTHORS  Ohnishi,M., Murata,T., Nakayama,K., Kuhara,S., Hattori,M.,
Kurokawa,K., Yasunaga,T., Yokoyama,K., Makino,K., Shinagawa,H. and
Hayashi,T.
TITLE    Comparative analysis of the whole set of rRNA operons between an
enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an
Escherichia coli K-12 strain MG1655
JOURNAL  Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
MEDLINE  20557356
PUBMED   11108008
REFERENCE
AUTHORS  Yokoyama,K., Makino,K., Kubota,Y., Watanabe,M., Kimura,S.,
Yutsudo,C.H., Kurokawa,K., Ishii,K., Hattori,M., Abe,H., Iida,T.,
Yamamoto,K., Hayashi,T., Yasunaga,T., Honda,T., Sasakawa,C. and
Shinagawa,H.
TITLE    Complete nucleotide sequence of the prophage VT1-Sakai carrying the
Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
O157:H7 strain derived from the Sakai outbreak
JOURNAL  Gene 258 (1-2), 127-139 (2000)
MEDLINE  20564182
PUBMED   11111050
REFERENCE
AUTHORS  Hayashi,T., Makino,K., Ohnishi,M., Kurokawa,K., Ishii,K.,
Yokoyama,K., Han,C.-G., Ohtsubo,E., Nakayama,K., Murata,T.,
Tanaka,M., Tobe,T., Iida,T., Takami,H., Honda,T., Sasakawa,C.,
Ogasawara,N., Yasunaga,T., Kuhara,S., Shiba,T., Hattori,M. and
Shinagawa,H.
TITLE    Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12
JOURNAL  DNA Res. 8 (1), 11-22 (2001)
MEDLINE  21156231
PUBMED   11258796
REFERENCE
AUTHORS  Hayashi,T., Ishii,K. and Shiba,T.
TITLE    Direct Submission
JOURNAL  Submitted (26-JUN-2000) Masahira Hattori, Kitasato Institute for

```

COMMENT

Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan
 (E-mail:hattori@genome.ls.kitasato-u.ac.jp,
 URL:http://genome.ls.kitasato-u.ac.jp/, Tel:81-42-778-8194,
 Fax:81-42-778-8193)
 genome project
 This work was done in collaboration with Tetsuya Hayashi, Makoto Ohnishi, Keisuke Nakayama (Miyazaki Medical College), Koza Makino, Ken Kurokawa, Katsushi Yokoyama, Masashi Tanaka, Takeshi Honda, Teruo Yasunaga, Hideo Shinagawa (Osaka University), Takahiro Murata (Shinshu University), Chang-Gyun Han, Eiichi Ohtsubo, Toru Tobe, Chihiro Sasakawa (University of Tokyo), Hideto Takami (Japan Marine Science and Technology Center), Naotake Ogasawara (Nara Institute of Science and Technology), Satoru Kuhara (Kyuushu University), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science.

FEATURES

```

source
1..307962
/organism="Escherichia coli O157:H7"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:83334"
72..653
/gene="Ecs3910"
72..653
/gene="Ecs3910"
/notes="similar to MDAB_ECOLI_gil1789406 percent identity 100 in 193 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="modulator of drug activity B"
/protein_id="BAB37333.1"
/db_xref="GI:13363383"
/translations="MSNLIITINGAKKFAHNSGQNLDTTEVADGTLRLDGLHVDVIRVA
DSPSKYSGGLVQGGKYMLSLTWNAPMEAFTEKDFHGVGDVGLVLPFHKQNFQFLGM
EPLPTIANDVIRKPDVPRYTEYRKHLEIFG"
684..998
/gene="Ecs3911"
684..998
/gene="Ecs3911"
/notes="similar to YGIN_ECOLI_gil1789407 percent identity 100 in 104 aa (Conserved in E.coli K-12)"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB37334.1"
/db_xref="GI:13363384"
/translations="MLTVIAETPRGQHHQAVLDQFAKIVPTVLKEGCHGYAPMV
DCAAGVSFQSMAPDSIVIEQWESIAHLEAHLQTPHKAYSEAVKGDVLEMNRILOP
GI"
complement(1043..1930)
/gene="Ecs3912"
complement(1043..1930)
/gene="Ecs3912"
/notes="probable transcriptional regulator similar to
regulators (Rpir family) e.g. [Bacillus subtilis]
gil18248807|emb|CAB93068.1| percent identity 25 in 236 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="BAB37335.1"
/db_xref="GI:13363385"
/translations="MIQOKTRVDIYGERFRFRASQLTFLGRAVASIYNEHREVVLDQ
TAMEIATNTSDATVIRAIQALGFAGLRDLKRLTEQWLGPALSSSEKMSSTVSNLTS
DVNTAIDFVLEGLHYTCNVLSEPNHALQAVALLVQARQVAIFGIGASGIADYTA
LRFNRICLPATALNRIGIQLAQDVLIMQAQSAHREGITLREARRLGIP
VILLTVALDSRFSKDSIVIHVPRGDEKGTPLHGTVLLCLEMIVMSVASVQRAVK
TIKRIIDFHRGLTKRNG"

```

gene	complement(1927. .2874) /gene="ECs3913" complement(1927. .2874) /gene="ECs3913" /note="probable ferrichrome-binding protein, similar to ferrichrome-binding proteins e.g. [Bacillus subtilis] gi 585132 sp P37580 RHUO_BACSU percent identity 27 in 220 aa" /codon_start=1 /evidence=not experimental /transl_table=11 /product="putative ferrichrome-binding protein" /protein_id="BAB37336.1" /db_xref="GI:13363386"	Best Local Similarity 99.0%; Pred. No. 1.4e-226; Matches 921; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
CDS	/translation="MRLFFSLILLSFFARATEPVQVETDILGRKVVTPAHPKRIVSL HDLDTIFLIEGVFPVASHGRTPDGFHRSFALLTGVDVFNSSIAFTADIDIE AIVAAKPOLIITEPTNPPIERLEKIATVSIHDKGAPBIYRKLAEITGQSQAI LERRVQAOINAKLIDSKITVSIQANQKINVMHSYHSLGRVLDAGFRFPPLIE STEGGRMDVSAERLPELDADVFVATWRDGTGGKQDELAAMEKVMPCWQCQFLTACS GRVLSREATNSPSFASIGLMAQIQSIAGRPLPEAK" complement(2885. .3934) /gene="ECs3914" complement(2885. .3934) /gene="ECs3914" /note="probable ferrichrome ABC transporter (permease) similar to ferrichrome ABC transporter (permease) e.g. [Bacillus subtilis] gi 1706795 sp P49937 FHUG_BACSU percent identity 33 in 319 aa" /codon_start=1 /evidence=not experimental /transl_table=11 /product="putative permease of ferrichrome ABC transporter" /protein_id="BAB37337.1" /db_xref="GI:13363387"	1 ATGGAACGACTAAACGCGATGTCGGTGTGTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60 213455 ATGGAACGACTAAACGCGATGTCGGTGTGTCGCAAGTAGTTGAATTTGGCTCTTTTACC 213514 61 GCGCGCCGACGAGCAGTACAGATGACGCTTTCGTCATCAGTCAGCAGGATCAAACTG 120 213515 GCGCGCCGACGAGCAGTACAGATGACGCTTTCGTCATCAGTCAGCAGGATCAAACTG 213574 121 GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACACGAGCAGTTCGCTGACCGGAA 180 213575 GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACACGAGCAGTTCGCTGACCGGAA 213634 181 GCGCGTAGAATTTACTACAGGGCTCCGTCGTATGCTTCATGAAGTGAGGATGTTTCA 240 213635 GCGCGTAGAATTTACTACAGGGCTCCGTCGTATGCTTCATGAAGTGAGGATGTTTCA 213694 241 GAGCAACTGATGCTTCAATACACCCCATCGGACGCTACGCAATGGCTGTTCTTCA 300 213695 GAGCAACTGATGCTTCAATACACCCCATCGGACGCTACGCAATGGCTGTTCTTCA 213754 301 ACTATGCGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGT 360 213755 ACTATGCGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGT 213814 361 TTGAGCGTCATCTGTTACCGGATTCAGCCCGGCTGATTCGCGACGCTGATTCGCGCGG 420 213815 TTGAGCGTCATCTGTTACCGGATTCAGCCCGGCTGATTCGCGACGCTGATTCGCGCGG 213874 421 GTGCGTCACTCGCGTCCGCGGCTTGACGAGTTTCAGGCTCTTTTCCGCGCTGTCGCGCGG 480 213875 GTGCGTCACTCGCGTCCGCGGCTTGACGAGTTTCAGGCTCTTTTCCGCGCTGTCGCGCGG 213934 481 ATGCCAATGTTGTTGTCGCGGGAAGATCTACACAAATACGGATACGGATACGGAA 540 213935 ATGCCAATGTTGTTGTCGCGGGAAGATCTACACAAATACGGATACGGATACGGAA 213994 541 CCGCGCGATTTGAGTAGTCAATCATGCTGCTGAATACAGCGTGGCGCGGCGCAATGAATTT 600 213995 CCGCGCGATTTGAGTAGTCAATCATGCTGCTGAATACAGCGTGGCGCGGCGCAATGAATTT 214054 601 GAATGATCGACCGGAGGAGTCTCGACTCGCTGATCCCAAGAGAGATTTGTGACT 660 214055 GAATGATCGACCGGAGGAGTCTCGACTCGCTGATCCCAAGAGAGATTTGTGACT 214114 661 AATGATCCGATGACGCTGTCGCTGCGTGGCGGCTGACGGGCTGCGCGGATCGCTACGTCGCG 720 214115 AATGATCCGATGACGCTGTCGCTGCGTGGCGTGGCTGACGGGCTGCGCGGATCGCTACGTCGCG 214174 721 CTGATGTCGGTGATCAACGAGATCAATCTGTCGGGAGCTCGAGATCTGTCGCGGCTTAC 780 214175 CTGATGTCGGTGATCAACGAGATCAATCTGTCGGGAGCTCGAGATCTGTCGCGGCTTAC 214234 781 CAGTCAGATCCAGCCCGCTTATCGCTTATATACCGAAAGATAGCTCGCGCTGAG 840 214235 CAGTCAGATCCAGCCCGCTTATCGCTTATATACCGAAAGATAGCTCGCGCTGAG 214294 841 GTACAGTCTGATCAACTCGCTGACGAGTATTTTGTGAGTTCGTTAAATTTGTTTCA 900 214295 GTACAGTCTGATCAACTCGCTGACGAGTATTTTGTGAGTTCGTTAAATTTGTTTCA 214354 901 GAGATGACGCGCGCGGGAAGAGATGA 930 214355 GAGATGACGCGCGCGGGAAGAGATGA 214384
gene	complement(4911. .5720) /gene="ECs3916" /note="probable ABC-type iron-siderophore transport system ATP-binding protein, similar to ABC-type iron-siderophore 98.5%; Score 915.6; DB 1; Length 307962;	RESULT 4 AE005552 LOCUS DEFINITION Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 171 of 290.


```

ACCESSION   AB005552
VERSION     AB005552.1
KEYWORDS    GI:12517866
SOURCE      Escherichia coli O157:H7 EDL933
ORGANISM    Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE   1 (bases 1 to 11593)
AUTHORS     Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
            Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
            Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,
            Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S.K.,
            Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,
            Welch, R.A. and Blattner, F.R.
TITLE       Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
JOURNAL     Nature 409 (6819), 529-533 (2001)
MEDLINE     21074935
PUBMED     11206551
REFERENCE   2 (bases 1 to 11593)
AUTHORS     Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D.,
            Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A.,
            Postfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L.,
            Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamou, S.K.,
            Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C.,
            Welch, R.A. and Blattner, F.R.
TITLE       Direct Submission
JOURNAL     Submitted (22-OCT-2000) Laboratory of Genetics, University of
            Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES    Location/Qualifiers
             1..11593
             /organism="Escherichia coli O157:H7 EDL933"
             /mol_type="genomic DNA"
             /strain="EDL933"
             /serotype="O157:H7"
             /db_xref="taxon:155864"
             /note="enterohaemorrhagic"
             88..1017
             /gene="gsea"
             /note="gsea"
             /gene="gsea"
             88..1017
             /gene="gsea"
             /function="putative regulator; Not classified"
             /note="putative transcriptional regulator LysR-type;
             Residues 1 to 309 of 309 are 99.67 pct identical to
             residues 1 to 309 of 309 from Escherichia coli K-12 Strain
             MG1655: B3243"
             /codon_start=1
             /transl_table=11
             /product="quorum sensing Escherichia coli regulator A"
             /protein_id="AAG58371.1"
             /db_xref="GI:12517867"
             /translation="MERLRKMSVFAKXFEFGSFTAAARQLQMSVSSISQTVSKLEDEL
             QVKLANRSTRIGLTAAGRIYVYQGRMLHEVDVQVHEOLYAFNNTPIGTLRIGCSSTM
             RQNLVAGXTAKMLKEVGLSVNLVGTIPADLIADGLDVIKVGALQDLSLFSRLGA
             RMVYCAKSVLTQGIPEKPADLSHSHWLSVYRPDNEFLIAPEGISRLIFQGRF
             VTNPMTLVRMTAGAGIAYVPLMWVINEIRNGLEILLPRYQSDRPVYALYTERDK
             LPLKQVYINSLTDYFVVGKLFQEMHGGRKEK"
             complement(1151..2596)
             /gene="tld"
             /note="tld"
             /gene="tld"
             complement(1151..2596)
             /function="phenotype; Not classified"
             /note="Residues 1 to 481 of 481 are 99.37 pct identical to
             residues 1 to 481 of 481 from Escherichia coli K-12 Strain
             MG1655: B3244"
             /codon_start=1
             /transl_table=11
             /product="suppresses inhibitory activity of CsrA"
             /protein_id="AAG58372.1"
             /db_xref="GI:12517868"
             /translation="MSLNLVSEQLIAANGKHKQDIFVLGQAEARRLDYGLYFQSSY
             HESWVLEDRIDKGSYNIDQGVGVAISGKGTGAYADQISLLALEQSAQAARTIARD
             SGDKGVQTLGAVEHSPFYTSVDPQLQMSREKDLIRRVVDKVAAREADKAKVQEVTSLSL
             GYVELIIVAAITDGTLAADVRPLVRLSVLIIEDGKRERGAAGGGGFGYEFPLADLD
             GEVRDAWAKEAVMALVNLASAAPAGTMRPVVAGPGLVLIHEAVGLGEGDFNR
             GTSVFGQVCELVASLCTVDDGTGTVRRGSVAIDDEGTPGOYNVLINGILKGMQYR
             DKNALMGWTPGTNGRRSVAHLPMRMTNTVLPKSTPOEIIESVEVGIYAPNG
             GGQVDITSGKFPSTSEAYLIENGKVPKVGATLIGSGIETMQQISMGVNDLKLNDG
             VVCGKGGSLPVGVCQPLKVDNLVGGTA"
             complement(2752..6552)
             /gene="Z4604"
             complement(2752..6552)
             /function="orf; Unknown function"
             /note="Residues 281 to 1266 of 1266 are 98.68 pct
             identical to residues 1 to 986 of 986 from Escherichia
             coli K-12 Strain MG1655: B3245"
             /codon_start=1
             /transl_table=11
             /protein_id="AAG58373.1"
             /db_xref="GI:12517869"
             /translation="MRRLPGILLTGAALVITAAALVSGRLALPHLDARPEILNKI
             ESAIGPVEASQSLASQNFPTLEAHDIAELKDGFEFSVKRVIALDNLVWQLLENR
             WQPRDTFWQLRFTNTPTITSGSDSLASHISDLFQDFHDFHDDLESEVSLTPSG
             QRAELAIQTLNDRRAEGLVSLTQGHVQVMQVMDLRDDEGLSLGVRWLQ
             ADDIDLPKLMQMDNIALETAQFSLGEMTIDKGVTDGVTGGVWLQGSASWLGKET
             HTLSVDNLTAHITRENPGQCFSPDTRITMDGKFPSPGALTIAWIPEQVGGKDKRS
             DELIRASNLLEAGLESGIRLAELKSPALGDVWRSTQPSKINTIALDIPLOADYTR
             FOASWDLANKWKPLPGRHFSCTLSGSENGLLTASMKQAKMPYETVFRPLEATD
             GQATISLNNKGFQDGRNIDVAKARHARGFRYLQFANDEFWLGISLTGSDGS
             QAWRYFENMLKDLVDYLSGAIQGGEDNATLVYGNPQLFPYKNEQGFVFLPLR
             NAKFAFQPDWALNTLIGDITNDGLMKTGVNIGVRSNLTAIPIDYSKELLI
             DADIKGKFAVGPYFDETLKDSLQALQELQDGDVNARLHLDIPLNGELVTAKEV
             TLNNSLFTKPLDSTLKNLSKGFSLNGLOQSEPLTASWNPQFLNDVFTKEGAKAYQ
             VAVLNGNQPAKTGVLPAVNEALSSVAMGKGVGLPYHAGATVYVELNGDLKNV
             SSHLPSPKAPAGEPLPVNVKYDGNLSFELTQAGADNHSRWLLGOKLLTLDRAIM
             AADSKTLPPLPQSGVELMPPMNGAEALFQKAESVGGASFPQHILTPMLS
             LGNQNNLSIVSQPTANGTQVEAQREINATLANNAPLANIKYLYNPSVTNGLI
             GDSPSSPFTTERINFRGMSDAQIRCAECWFQKGRFIDSDITISGNTLTNGLI
             DTGFSLTADGEMINNGERTSLKGLRQKIDAAAEFFGVTPTRQSSFNVDYDLH
             WRKAPWDEATLNGIHTQLGKEITTEINTGHAGOLLRLLSVDALMKRLRFPDFTF
             GEGFYDSIRSTAWIKDGMVHTDDTLVDGLEADIAMKSGSVNLVRDLNNEAVVAPIS
             ATVGVAAPVAVNPIVGAAPFAASKVLGPLNSKVSILRYHISGLPDLDPQINEVLQRPK
             EKAQ"
             complement(6620..8107)
             /gene="cafa"
             /note="synonym: Z4605"
             complement(6620..8107)
             /gene="cafa"
             /function="structural component; Cell division"
             /note="Residues 1 to 495 of 495 are 100.00 pct identical
             to residues 1 to 495 of 495 from Escherichia coli K-12
             Strain MG1655: B3247"
             /codon_start=1
             /transl_table=11
             /product="bundles of cytoplasmic filaments"
             /protein_id="AAG58374.1"
             /db_xref="GI:12517870"
             /translation="MRKGINMTAELNVNTPSETRVAYIDGGILQBIHIEREARRGIV
             GNTRYGRSVRLPQMAAFVDGLDKAAFLHASDIMPTCEVAGEKQFTVRDISEL
             VROGDLVQVWVDPLGTGARLTDTITLPSRYLPMGASHVGSORIESERRRL
             KTVVAVCDQGGFIIRTAEGVGEALSDAAYLKRVTWKMRKQRTQRYQLYGE
             LALAQRVLRDPADELDRVDSRLTYEALLFTSEYIPEMTSKLEHYTGRQPIFLF
             DVNEIQRALERKVELSGYLIDQTEAMTVINDTNGFVGRNDDDTIFNINBAT
             QATARGLRNLGGIIIDIFDMNEDHRRVLHLSLEQALSKDRVTSYNGSGLGV
             EMTRKTRTSIEIHLNCEGTCHGTQVETVTCVEINREIVRVHDSRPLVYA
             SPVABALAGESHSLAEVSIFFVGKQVQIEPLVNEQGFVVMV"
             complement(8079..8672)
             /gene="yhdE"
             /note="synonym: Z4606"
             complement(8079..8672)
             /gene="yhdE"
             /function="orf; Unknown function"
             /note="Residues 1 to 197 of 197 are 98.47 pct identical to
             residues 1 to 197 of 197 from Escherichia coli K-12 Strain

```


Schwartz, D.C. and Blattner, F.R.
Direct Submission
Submitted (13-JUN-2002) Genetics Laboratory, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 292497
/organism="Shigella flexneri 2a str. 2457T"
/mol_type="genomic DNA"
/strain="2457T"
/serotype="2a"
/db_xref="taxon:198215"
98..1169
/gene="ygjJ"
/locus_tag="S3326"
/note="frameshift"
pseudo
1186..3537
/gene="ygjK"
/locus_tag="S3327"
1186..3537
/gene="ygjK"
/locus_tag="S3327"
/function="putative enzyme; Not classified"
/note="residues 1 to 783 of 783 are 95.65 pct identical to residues 1 to 783 of 783 from Escherichia coli K-12 : B3080"
/codon_start=1
/transl_table=11
/product="putative isomerase"
/protein_id="AAP18405.1"
/db_xref="GI:30042682"
/translation="MKITLIPVTCALLISFAHAAADNKNVINTGAPQWKMYYDQHQFNFFDLGAWHLLPDGNTMGPGFVALITEEYINFMASNFRLTVWQD GKXNFTLEAYIPGALVQKLTAKDQVEMTLRFPATPRTSLLETKIISNKPLDLVWQD ELLEKLEAGSGPLDKTAGEPDYQKISATRDGLKVTGKVAZWDLLTSGSEY OVHSLPVQTEINGNFTSKAHINGSTLTYYTSHLLTAQVSKQEMOIRDILAPAF YLTASQWEEYLYKGLNPDATPEQTRVAVKAIETLNGNRSPCGAVYNTVPSVT GRWFGNQPDWTKWQFAMAFNPDIKENIRAVFSQIQPGSVRPQDVGFPDIL IAWNLSPEGGQGNWNETKPSLAASVMEVYNTQDKAWLAEMYPLVAYHDWDL RNRDHNGVPGATRDKAHNTSEGLFTVKGDKBEQSGNLNRYARVVEKQYDS LEIPQAASWESGDDAAVFGFIDKEQLDKYVANGKRSQDWTVKFAENRQDQTLG YSLQESVDQVMSYDNHYLAEMATILGKPEAKRYQLAQALADYINTCWFQPTVO FYDVRIDKPLANGCAGKPIVERGKPGESPLFNGAQTQANADPVKMLDPKEFN TVPLGTALTPARGADLYWRGVWQDFWFGFKGMEYGYRDDALKLATDFFRAHX GUTAGPQIENFTNLTGACQGFAPNFSWGAHLIYMDINDFFRQ"
3550..3632
/note="REP (repetitive extragenic palindromic) element, contains 2 REP sequences"
3763..5781
/gene="ygjL"
/locus_tag="S3328"
3763..5781
/gene="ygjL"
/locus_tag="S3328"
/note="residues 1 to 672 of 672 are 95.23 pct identical to residues 1 to 672 of 672 from Escherichia coli K-12 : B3081"
/codon_start=1
/transl_table=11
/product="putative NADPH dehydrogenase"
/protein_id="AAP18406.1"
/db_xref="GI:30042683"
/translation="MSYPSLPAFLDFTLLKNRVLMGSMHTGLEEYPDGAERLAIFY AERARHVALIVSGGIAPDLTGVMGGAMLDASQIPHHSTITEAVHOGGKIALOI LHTGRYSQPHLVASALQAPINRVFPHLTHEELQLIDDFARCAQARAGVDGVE VNGSGYLNEFTLTKQRSDQWGDYRNRFVAVVTVRVRVGNDFIIIVYLSM LDVLEGGFFATFELAQAEAGATINTGIGHEARIPTIATVPFGASVWTRKL KGHSLPLVTRNRPDQVADILSRGADIVSMARPLADAEALLSKAQSRADEINT CIGCNGQDLQIFVGKVTSLVNPRACTKMPILPAVOKKNLAVVGAGPAGLAFAIN AAARGHVTLPFAHSEIGQFENI AKQIPCKEFTYETLAVYRMIEVTGTVLKNHTVT AQLQAPETILLASIVRPPIIDGIDHPKVLSDYLDVRDKAPVGNKVAILGCGGIGF DRAITUSQESTSIARFPCNWDGIDSLOQAGLSFGQMCIPKSPQIVMLQKAS KPGQGLGKTGWIHRTLLLSRGVRKMPGVSYQKIDDDGLVHVGINGETQVLAVDNVIC

gene
CDS
gene
CDS
repeat_region
gene
CDS
gene
CDS

AGOPNFALAQPLIDSGKTVHLIGGCDVAVMELDARRAIAQGTRLALEI"
complement(5826..6242)
/gene="ygjM"
/locus_tag="S3329"
complement(5826..6242)
/gene="ygjM"
/locus_tag="S3329"
/note="residues 1 to 138 of 138 are 90.57 pct identical to residues 1 to 138 of 138 from Escherichia coli K-12 : B3082"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18407.1"
/db_xref="GI:30042684"
/translation="MIAIDILQAGEKLTAVAPLAGIONEQYTOALELVVDHLLND PENPLDLVCAKITANWESAPFAENPMAQMPGGIATVIRLMDQVGLTSLDLPEIG SKMVSRLVSGKKRLTLESHAKPLATRGISPAULFI"
complement(6239..6553)
/gene="ygjN"
/locus_tag="S3330"
complement(6239..6553)
/gene="ygjN"
/locus_tag="S3330"
/note="residues 1 to 104 of 104 are 100.00 pct identical to residues 1 to 104 of 104 from Escherichia coli K-12 : B3083"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18408.1"
/db_xref="GI:30042685"
/translation="MHLITQKALKDAEKYPOHKTELVALGNTIAKGYFKKPSLKAV FPLSDNPKYLDKHVFNUGNELRVVAMVFFESQKCYIREVMTKEYDYDFTAVHRTKG KK"
6736..6822
/note="REP (repetitive extragenic palindromic) element, contains 2 REP sequences"
complement(6838..7974)
/gene="ygjO"
/locus_tag="S3331"
complement(6838..7974)
/gene="ygjO"
/locus_tag="S3331"
/functions="putative enzyme; Not classified"
/note="residues 1 to 378 of 378 are 95.76 pct identical to residues 1 to 388 of 388 from Escherichia coli K-12 : B3084"
/codon_start=1
/transl_table=11
/product="putative enzyme"
/protein_id="AAP18409.1"
/db_xref="GI:30042686"
/translation="MSHLNDFRSLTQRPFPATDDVNPLOAWAADADEYLLQQLDDTEI LGSVLINDAFGALSALAEHKPSYSGDSVISELATRENRLNGIDESSVKFLDSTAD YPOQGVLLIKVPKTLALLLEQOLARLKVVSPTRIIAGAKARDIHTSTLEPEKVLG PTTTTLANKARLINCTFNEPELADAPQVSKLEGTDMTINHNAVRSRTGLDIGAR FFMQHPENLEGEIVLGCNGVIGLTLDKNPQAKVVFVDESFMVAVASRLNVEINM PEALDRSEFMINNALSGVBEFFNFVAVLNCNPPPHQHQHALTDNVAVEMFHYARRCKING ELYIVANRHLDDYFHLKLLKIFGNCCTTIATNNKFVYLKAVKLGRR"
8023..8562
/gene="ygjP"
/locus_tag="S3332"
8023..8562
/gene="ygjP"
/locus_tag="S3332"
/note="residues 1 to 179 of 179 are 98.32 pct identical to residues 1 to 179 of 179 from Escherichia coli K-12 : B3085"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18410.1"


```

/note="Residues 1 to 309 of 309 are 99 pct identical to
residues 1 to 309 of a 309 aa protein from Escherichia
coli O157:H7 ref: NP_312143.1"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator LYSR-type"
/protein_id="AA04746.1"
/db_xref="GI:24053720"
/translation="MERLKRMSVFAKVVEFGSFTAAARQLQMSVSSISQTVSKLEDEL
QYKLLNRTSITGLTEAGRIYQGRMLHEVQDVEHQLYAFNNTPTGLRIGCSSTN
AGNVLATKMLKEYPGLSANLVGTI PAPDIIADGLDVIWIRGALQDSLSFSRRLGA
QMNVYCAKSYLTQYGIPEKPADLSHSLWLSYVRPDNEFELIAPEGISTRLIPQGRF
VTNDPMTLVRLTTCAGIAYVPLMWVINEINRGELEILLPRYQSDPRVVALYTEKDK
LPLKVQVINSITDYFVVGKLPQEHGGRGKEK"
complement(1285..2730)
/gene="tldD"
/locus_tag="SF3283"
/complement(1285..2730)
/gene="tldD"
/locus_tag="SF3283"
/note="Residues 1 to 481 of 481 are 100 pct identical to
residues 1 to 481 of a 481 aa protein from Escherichia
coli K12 ref: NP_417711.1"
/codon_start=1
/transl_table=11
/product="suppresses inhibitory activity of CsrA"
/protein_id="AA04747.1"
/db_xref="GI:24053721"
/translation="MSLNLVSEQLLAANGLKHQDLFALQLAERLLDYGLYQSSY
HESVLEDRITKDSYNDQGVRAISGEKTFAYADQISLLALEQSAQAARTVRD
SDGVRQITAGVSHFTYSDVPLQMSREKLDLIRRVKVARADKRVQEVTAISL
GYVELIWAADTGLTAAVDRPLVLSVSVLVEEDKRGKRGASGGGGRFGYFFFLADIL
GEBVDAKAVARLNLNSAAPAAGTVPVILGAGPGLLHEAVGHLEGGDFNRR
GTPVSGGVASELACTVDDGTMDRGRSVAIDDEGTQGVNVLLENGILKGYMG
DKLNRMGMTFTNGRRESYAHLPMPNTYMLPGKSTPQEIIESVEYGIYAPNFG
GQVDITGKFTSTSEAYLIENGKVTTPVTKATILGSGIETMQQISVGNGLKNDG
VGCQEGQSLPVGQGTLPKVDNLTVGGTA"
complement(2886..6686)
/gene="yhdR"
/locus_tag="SF3284"
/complement(2886..6686)
/gene="yhdR"
/locus_tag="SF3284"
/note="Residues 1 to 1266 of 1266 are 99 pct identical to
residues 1 to 1266 of a 1266 aa protein from Escherichia
coli K12 ref: NP_417712.1"
/codon_start=1
/transl_table=11
/product="uncharacterized membrane protein"
/protein_id="AA04748.1"
/db_xref="GI:24053722"
/translation="WRRILPGILLTGAALVIAALLVSLRILALPHLNAWRPEILNKI
ESATGMPVEASOLSWONFGPTLEAHDIRAEIKDGFPSVKRVTLADVWQSLHMR
WQFDLTFWQLEFRNTPTITSGSDSDSLEASHISPLRQPDHDLRSEVSLTPSG
QRAELAIQLTLWLNPRHRAEGLVLSLGTQGHGVQVRMDLRDDEGLLNGRVLQ
ADDIDKPLGWQMNQIALETAQFSLGEMWTIDKGVDTGGVWMLQKGGASWLGEKQT
HTLVNDNTAHITRENQWQFSIPDTRI.TMDGKWPSPALTLAMIPEQDVGNKKRS
DELRASNLLEAGLEGRPLAAKLSPALGDVWVRSTQPSKINTLALDIPQAADKTR
FOASWDLAKWKLLPGAHEFSGTSLSGVENGLITASMKQAKMPEYTVFRAPLEIAD
GQATISVNNDKFOLDGRNIDVKAHARGCPRYLOPANDPEMLGILAGISTDDGS
GOATRYPENLMDGLDVLVLSGAIQGEADNATLVYGGNPQLPYKHNQEPFVLVPLR
NAKFAQPDWPAITNLDIELDFINDGLWKKTDGCVNLGVRASNLTAVIDPDSKXLLI
DADIKPGKAVGPFYDETPDKSLGATQELQDGDVNARLHLDIPLNGELVITKAGEV
TURNSLFTKPLDLSIKNLGKGFSPINDGLSEPLTASFNQPLNDFSTKEGAKYC
VAVNLNGWQPAKTGVLEAVNEALSGSVANDGVKGI DLPYHAGATYVNLGLNNV
SGLHPSLPKAPGELVAVNVKVDGNLNSFELTQAGADNHFNSRWLLQKLTLDRAIV
AASDKTLPPLPQSGVANGLMPMNGAERLALPOKGAESVGAASFPOHITLRTPLMS
LNCQOWNLSIYQSGVELTVEAQRINATLAVRNNAPLANIKVLYNPSVAKTR
GSTPSSPPTTERINFRGNPDQIRCTECFWGQKFGRIISDITISGDTITLNGLI
DTGFRSLTAGDGVNPNPGRNRTSLKURQKIDAAEFFGVITPIROSSENVDDH
WRKAPQWQDEATLNGIITHTQKGBEITINTGHACQLELLSVDMKRFDFRDTF
GGGFYDSTRSTAWIKDGMHTDDTLVDGLEADIAKMSVNLVRDPLNNEAVAPEIS
ATVGVAAAFVNPVIGAAVFAASKVLGLPWSKVSILRYHSIGPLDDPDQINEVLRQPRK

```

```

EKAQ"
complement(6754..8241)
/gene="cafa"
/locus_tag="SF3285"
/complement(6754..8241)
/gene="cafa"
/locus_tag="SF3285"
/note="Residues 1 to 495 of 495 are 99 pct identical to
residues 1 to 495 of a 495 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289814.1"
/codon_start=1
/transl_table=11
/product="bundles of cytoplasmic filaments"
/protein_id="AA04749.1"
/db_xref="GI:24053723"
/translation="MRKGINMTAELLVNTPTSRVAYIDGILQBIHIERARRGIV
GNIYKGRVSRVLPFGMAAFVDIGLKAALFHASDIMPHTECVAGEQKQFTVRDISEL
VROQDLVQVWVDPLGTGKARLTITDITLPSRYLVFPMGASHVGVQSRLESERREL
KVVAVYCDQGGFIIRTAAGVGGEAEALDAAYLKRVWTKVMERKKRQTRYQLYGE
LALAQVLRDEADAELDRVDSRLTYEALLETSEYIPEMTSKLHYTGROPIFDLF
DVNEIQORALERKVELKSGYLIIDOTEAMTVDINTGAFVGHENLDDTIFNINIEAT
QAIARQRLNGLGIIILDFIDNNSDHRRVRLHSLQALSKDRVKTSTVNGPSALGIV
EMTKRTRRESIEHVLNCEGPTCHGRGTGVKTIVTCYEIMREIVRVVHHAYDSDFLVIA
SPAAEALKGEESSLAEBEIVFGVKQVQIEPLYNQEPDVMVM"
complement(8213..8806)
/gene="yhdE"
/locus_tag="SF3286"
/complement(8213..8806)
/gene="yhdE"
/locus_tag="SF3286"
/note="Residues 1 to 197 of 197 are 97 pct identical to
residues 1 to 217 of a 217 aa protein from Escherichia
coli gb: AA58051.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA04750.1"
/db_xref="GI:24053724"
/translation="MTSLYLSGSPRRQELLAQLVTFERIVTGIEEQROPBSAQOY
VVRAREKAQAGVATQADLPVLGADIIVILNGEVLKPRDAEHAQMRLKLSGQTHQ
VMTAVALASQHILDCLVVDITFRLTDEIDAGYVASGEPLDKAGAYGIGLGGCFV
RKINGSYHAVVGLPLVETVELLSNFALREKDKHDG"
complement(8815..9303)
/gene="mreD"
/locus_tag="SF3287"
/complement(8815..9303)
/gene="mreD"
/locus_tag="SF3287"
/note="Residues 1 to 162 of 162 are 98 pct identical to
residues 1 to 162 of a 162 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289816.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA04751.1"
/db_xref="GI:24053725"
/translation="MASYRSQGRWVNLFLIALLQIWPDPNLVFRPNWVLLILL
YTLIALPVRNVGTGVMGAILDLISGLVGVLAISYIALVALKYQLRNLALWQ
QALVWMLLSLVVDIIVFAEFLVINTVSPRPEVFWSSVAVNGVLPWFIFLLMRKVRQOFA
VQ"
complement(9303..10406)
/gene="mreC"
/locus_tag="SF3288"
/complement(9303..10406)
/gene="mreC"
/locus_tag="SF3288"
/note="Residues 1 to 367 of 367 are 99 pct identical to
residues 1 to 367 of a 367 aa protein from Escherichia
coli K12 ref: NP_417716.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AA04752.1"

```

```
/db_xref="GI:24053726"
/translation="MKPIFSRSPS:QIRLILAVLVALGIIIDSELGTFESQIRTYMDT
AVSPFYFVSNARELLDGIQSLARLDQLEENRALRQELLKNSLMLQYKQENA
RURELLGSLRQDQRMVITQVISTVNDPFDSELDVDRGVSNGVIEGQVSDKGVGQ
VVAVALKLSRVLLI CDATHALPIQLRNDIRVIAAGNCTDDDLHLHPANTDIRVGD
VLVTGLGGRFPEGVAVVSSVKLDTQAYTIOARPTAGLQRLYLLLLWADNRNG
ANPMTPPEVHRVAMERLMMQPVLPSPDAMPKLPPEPATGIAOPTPOCPATGNAATA
PVATPCPAARSPORATPQSGAOPPARAPGQ"
complement (10472..11590)
/locus_tag="SF3289"
/locus_tag="SF3289"
Query Match 98.1%; Score 912.4; DB 1; Length 11734;
Best Local Similarity 98.8%; Pred. No. 7.8e-226;
Matches 919; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 ATGGAACGACTAAACGCGATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 60
DB 228 ATGGAACGACTAAACGCGATGCGGTGTTGCCAAGTAGTGAATTTGGCTCTTTTACC 287
QY 61 GCGCGCCGACAGCAGCTACAGATGAGCGTTCCTCCATCAGTCAGACGCTATCAAACTG 120
DB 288 GCGCGCCGACAGCAGCTACAGATGAGCGTTCCTCCATCAGTCAGACGCTATCAAACTG 347
QY 121 GAAGATGAGTTCGAGTAAAGCTGTTAAACCGTAGCAGACGAGATTGGCTGACCGAA 180
DB 348 GAAGATGAGTTCGAGTAAAGCTGTTAAACCGTAGCAGACGAGATTGGCTGACCGAA 407
QY 181 GCGGTGAGATTTACTACGAGGCTGCCCTGCTATGCTTCATGAAGTCAGAGATGTTTCA 240
DB 408 GCGGTGAGATTTACTACGAGGCTGCCCTGCTATGCTTCATGAAGTCAGAGATGTTTCA 467
QY 241 GAGCAACTGTATGCTTCAATTAACCCCATCGGAGCGCTACGATGCGTGTGTTCTTCA 300
DB 468 GAGCAACTGTATGCTTCAATTAACCCCATCGGAGCGCTACGATGCGTGTGTTCTTCA 527
QY 301 ACTATGGCACAAATGTTCTCCGCGGTGACAGCCAAATCTCAAGAAATACCCAGGT 360
DB 528 ACTATGGCACAAATGTTCTCCGCGGTGACAGCCAAATCTCAAGAAATACCCAGGT 587
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCACGCCCCCGACCTGATTCGCGAGCGTCTGAT 420
DB 588 TTGAGCGCAATCTGTTACCGGAATTCACGCCCCCGACCTGATTCGCGAGCGTCTGAT 647
QY 421 GTGGTATCGCGTCCGCGCTGAGGATTCAGGATTCAGGCTGTTTCCGCGCTCGGCGCG 480
DB 648 GTGGTATCGCGTCCGCGCTGAGGATTCAGGATTCAGGCTGTTTCCGCGCTCGGCGCG 707
QY 481 ATGCCAATGTTGTCGCGCGGAAAGCTATCTCACAAATACGGATACCGGAAAAA 540
DB 708 ATGCCAATGTTGTCGCGCGGAAAGCTATCTCACAAATACGGATACCGGAAAAA 767
QY 541 CCGCGCGATTTGAGTAGTCAATTCATGCTGTTGAATACAGGTCGCGCGCGGACAAATTT 600
DB 768 CCGCGCGATTTGAGTAGTCAATTCATGCTGTTGAATACAGGTCGCGCGCGGACAAATTT 827
QY 601 GAAGTATCGACCGGAGGAGTCTCGACTCCCTGATCCCAAGAGGAGATTTGTGACT 660
DB 828 GAAGTATCGACCGGAGGAGTCTCGACTCCCTGATCCCAAGAGGAGATTTGTGACT 887
QY 661 AATGATCCGATGACGCTGCTGCGCTGGCTGACGCGGCGGTCGCGGATCGCCTACGTCGC 720
DB 888 AATGATCCGATGACGCTGCTGCGCTGGCTGACGCGGCGGTCGCGGATCGCCTACGTCGC 947
QY 721 CTGATGCGGTGATCAACGAGATCAATCTGCGGAGCTGAGATCCTCTCCGCGGTTAC 780
DB 948 CTGATGCGGTGATCAACGAGATCAATCTGCGGAGCTGAGATCCTCTCCGCGGTTAT 1007
QY 781 CAGTCAGATCCAGCGCCGCTTTATCGGTTATATACCGGAAAAAGATAGCTCCGCTGAAG 840
DB 1008 CAGTCAGATCCAGCGCCGCTTTATCGGTTATATACCGGAAAAAGATAGCTCCGCTGAAG 1067
QY 841 GTACAGGTCGTGATCAACTCGCTGACGAGTATTTTGTGAGTCCGTAATTTGTTTCAG 900
```

```
Db 1068 GTACAGTCTGATCAACTCGCTGACCGATATTTTGTGAGTCCGTAATTTGTTTCAG 1127
QY 901 GAGATCAGCGGCGCGGAAAGAGAGTAA 930
DB 1128 GAGATCAGCGGCGCGGAAAGAGAGTAA 1157
RESULT 7
LOCUS AE016767 300359 bp DNA linear BCT 24-DEC-2002
DEFINITION Escherichia coli CFT073 section 13 of 18 of the complete genome.
ACCESSION AE016767 AE014075
VERSION AE016767.1 GI:26110054
KEYWORDS
SOURCE Escherichia coli CFT073
ORGANISM Escherichia coli CFT073
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 300359)
AUTHORS Welch,R.A., Burkland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence
of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)
12471157
REFERENCE 2 (bases 1 to 300359)
AUTHORS Welch,R.A., Burkland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
FEATURES
Location/Qualifiers
1: 300359
/organism="Escherichia coli CFT073"
/mol_type="genomic DNA"
/strain="CFT073"
/db_xref="taxon:199310"
188..547
/genes="yqiC"
/locus_tag="c3796"
188..547
/genes="yqiC"
/locus_tag="c3796"
/notes="Escherichia coli K-12 ortholog: b3042; Escherichia
coli O157:H7 ortholog: z4400"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiC"
/protein_id="AAN82241.1"
/db_xref="GI:26110055"
/transl_table=11
/translation="MTYPTTMAALNTYQYRETTWIDPKKIEQIAQVHSMKPIRGE
FEDVKKIKRQLQALQRLDLVSRBEFDVQTVLLRTREKLALLQRISELEARNP
VAELQSPPAIPVDKAE"
complement (582..788)
/genes="gls"
/locus_tag="c3797"
/locus_tag="c3797"
complement (582..788)
/genes="gls"
/locus_tag="c3797"
/function="putative enzyme; Macromolecule synthesis,
modification: Polysaccharides - (cytoplasmic)"
/notes="Escherichia coli K-12 ortholog: b3049; Escherichia
coli O157:H7 ortholog: z4401"
/codon_start=1
/transl_table=11
/product="Glycogen synthesis protein gls"
/protein_id="AAN82242.1"
/db_xref="GI:26110056"
```

```

/translation="MLMDHSLNSLNNFDFLARSFARMVHAEGRPVDFILAVTGNVDEBHR
TWFCARWAYVCOQMMQTFRELELEH"
gene
1052..1681
/gene="yqiJ"
/locus tag="c3798"
1052..1681
/gene="yqiJ"
/locus tag="c3798"
/function="putative"
/note="Escherichia coli K-12 ortholog: b3050; Escherichia
coli O157:H7 ortholog: z4402"
/codon start=1
/transl table=1
/product="Hypothetical protein yqiJ"
/protein_id="AAN82243.1"
/db_xref="GI:26110057"
/translation="MILFADYNTPYFAISFVLLIGLELIFALICGHMISGALDAHL
HYDSITGHISQALHYSGLIPALVVLCLAGFLIGLLQACVTLWQSPPLNLF
VVPVSLITFAVHTGKWAPMIPROHSAITEEVVGSMALITGHQATSGNPGCEGK
LTDQFGOIHVYLLLEPESGKFTKGDKVLIICRLSATRYLAENPNWSQL"
1708..3369
/gene="yqkK"
/locus tag="c3799"
1708..3369
/gene="yqkK"
/locus tag="c3799"
/function="putative membrane"
/note="Escherichia coli K-12 ortholog: b3051; Escherichia
coli O157:H7 ortholog: z4403"
/codon start=1
/transl table=1
/product="Hypothetical protein yqkK"
/protein_id="AAN82244.1"
/db_xref="GI:26110058"
/translation="MDDIVNSVFSWMTAIIAVCILFIIGIIPARLYRASABQAFVR
TGLSGQKVMGSGAIVMPIPHEIIPINNTLKLVEKSTIDSITKDRMVDVVVAF
VRVPSVEGATLQATGILGTSPEDLMIVEDKFDVALRTAAQMTMELQDTRRF
VQGVNTVAEDLNSKGLSEVSLTNFQTSKEHFNPNFADEGLTKLTQETERRR
ERNVEQDVEAVREKRDALSRKLETEQOEAVTEQEOQVKTAEONAKIAAPEA
ERREAEQTRILAEQITQETIDREQAVRSKVEAREVRKIEIQOQVTEIANOTKS
IAIAKQSQSQAEARLANLAELAVSAQQNVETTRCTAEDRAKQVALIAAQDAETK
AVETLVKAEKEAENMQAAIVELAEATRKKGLAEARALNDAINVLSDEQTSLK
FKLALLQALFAVIEKSPKMAIDGIKIIQVDNRRGAGTGDTANGVGGGNLAEQAL
SAALSYRTQAPLDSLNETGVSGGSAAALTSLSSTTPVEEKAE"
complement(3788..5221)
/locus tag="c3800"
complement(3788..5221)
/locus tag="c3800"
/EC_number="2.7.1.-"
/function="putative"
/note="Escherichia coli K-12 ortholog: b3052; Escherichia
coli O157:H7 ortholog: z4405"
/codon start=1
/transl table=1
/product="ADP-heptose synthase"
/protein_id="AAN82245.1"
/db_xref="GI:26110059"
/translation="MKVTLPEFERAGVWVDMLDRYVWYGTSPTRISPEAPVWVKVN
TIERPGCAANVANNISAGANARLVGLTGDDAARLAKSLADNVNKCDFVSPVTHP
TITVLRKVRNQOILRDPEGPEGVDPPQHERINQALSSICALVLSYAKGALASV
QOQTLQARAGAPVILIDPKGTDEVEGATLLTNISEFAVVGKCKTEELIVERGKK
LIADYELSAVVTRSEQSSILQPKAPLHMQAQEVDVTGAGDTIVGLVALIAVLA
KNSUEAEKCFANAAAGVNGKLGSTVSPKIELENAVGRADTGFVMTSEELKSLVAA
ARKKGEKVTWNGVFDILHAGHSYLANRKLGRILAVNSDASTRKLGKSRPVPN
LEQRMIVIGALEADVWVSEEDTPQRLIAGILPDLLVKGDKGYKPEIAGSKEVWANG
GEVILVNPEDGCSYTNIIKKIQIDKKG"
complement(5269..8220)
/gene="glnE"
/locus tag="c3801"
complement(5269..8220)
/gene="glnE"
/locus tag="c3801"
/EC_number="2.7.7.42"

```

```

QY 61 GCGCGCCGACAGCTACAGATGAGCGTTTCCTCATCAGTCAGACGGTATCAAACTG 120
Db 182674 GCGCGCCGACAGCTACAGATGAGCGTTTCCTCATCAGTCAGACGGTATCAAACTG 182733
QY 121 GAAGTAGTTCGAGTAAAGCTGTTAAACCTAGCAGCAGCAGCAGTGGCTGACCAA 180
Db 182734 GAAGTAGTTCGAGTAAAGCTGTTAAACCTAGCAGCAGCAGCAGTGGCTGACCAA 182793
QY 181 GCGGTAGAAATTTACTACAGGCGTCCGCTGCTATGCTTTCATGAAGTCAGAGTCTTTCAT 240
Db 182794 GCGGTAGAAATTTACTACAGGCGTCCGCTGCTATGCTTTCATGAAGTCAGAGTCTTTCAT 182853
QY 241 GAGCACTGTATGCTTCAATACACCCCTCGGAGCGCTACGATGGCTGTCTTCTCA 300
Db 182854 GAGCACTGTATGCTTCAATACACCCCTCGGAGCGCTACGATGGCTGTCTTCTCA 182913
QY 301 ACTATGGCACAAATGTTCTCGCGGGCTGACAGCAGCAAAATGCTGAAAGATACCCAGGT 360
Db 182914 ACTATGGCACAAATGTTCTCGCGGGCTGACAGCAGCAAAATGCTGAAAGATACCCAGGT 182973
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGGACCTGATTCGCGAGGTCGTGAT 420
Db 182974 TTGAGCGTCAATCTGTTACCGGAATTCGCGCGGACCTGATTCGCGAGGTCGTGAT 183033
QY 421 GTGGTATCCGGTGGCGCGTTCGAGGATTCAGCCTGTTTCCGCGCTGGCGCG 480
Db 183034 GTGGTATCCGGTGGCGCGTTCGAGGATTCAGCCTGTTTCCGCGCTGGCGCG 183093
QY 481 ATGCCAATGGTGTGCGCGCGGAAAGCTATCTCACAAATACGGATACCGGAAAAA 540
Db 183094 ATGCCAATGGTGTGCGCGCGGAAAGCTATCTCACAAATACGGATACCGGAAAAA 183153
QY 541 CCGCGCGATTTGAGTAGTCAATTCATGCTGCTTGAATACAGGTCGCGCGGACAAATGAT 600
Db 183154 CCGCGCGATTTGAGTAGTCAATTCATGCTGCTTGAATACAGGTCGCGCGGACAAATGAT 183213
QY 601 GAAGTATCGCACCGGAGGATCTCGACTCCGCTGATCCCAAGGAGATTTGTGACT 660
Db 183214 GAAGTATCGCACCGGAGGATCTCGACTCCGCTGATCCCAAGGAGATTTGTGACT 183273
QY 661 AATGATCCGATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 183274 AATGATCCGATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183333
QY 721 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 183334 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183393
QY 781 CAGTCAGATCCAGCGCGCTTTATCGTTATATACGAAAGATGATGCTGCTGCTGCTGCTGCT 840
Db 183394 CAGTCAGATCCAGCGCGCTTTATCGTTATATACGAAAGATGATGCTGCTGCTGCTGCTGCT 183453
QY 841 GTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 183454 GTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 183513
QY 901 GAGATGACGCGCGCGGGAAGAGAGTAA 930
Db 183514 GAGATGACGCGCGCGGGAAGAGAGTAA 183543

```

```

RESULT 8
AC020970
LOCUS AC020970 212936 bp DNA linear HTG 10-FEB-2000
DEFINITION Mus musculus clone RP23-252M21, WORKING DRAFT SEQUENCE, 144
unordered pieces.
ACCESSION AC020970
VERSION AC020970.1 GI:6691260
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.

```

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

1 (bases 1 to 212936)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 212936)
DOE Joint Genome Institute.
Direct Submission
Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>
-----Summary Statistics
Estimated insert size: 212936; sum-of-contigs estimation
Estimated insert size: 157000; pulse field gel estimation
Quality coverage: 4.20x in Q20 bases; pulse field gel estimation
Quality coverage: 3.10x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 144 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1159: contig of 1159 bp in length
* gap of unknown length
* 1160 2363: contig of 1204 bp in length
* gap of unknown length
* 2364 3604: contig of 1241 bp in length
* gap of unknown length
* 3605 5188: contig of 1584 bp in length
* gap of unknown length
* 5189 6372: contig of 1184 bp in length
* gap of unknown length
* 6373 7737: contig of 1365 bp in length
* gap of unknown length
* 7738 9783: contig of 2046 bp in length
* gap of unknown length
* 9784 11046: contig of 1263 bp in length
* gap of unknown length
* 11047 12223: contig of 1177 bp in length
* gap of unknown length
* 12224 13492: contig of 1269 bp in length
* gap of unknown length
* 13493 14669: contig of 1177 bp in length
* gap of unknown length
* 14670 15886: contig of 1217 bp in length
* gap of unknown length
* 15887 17089: contig of 1203 bp in length
* gap of unknown length
* 17090 18308: contig of 1219 bp in length
* gap of unknown length
* 18309 19518: contig of 1210 bp in length
* gap of unknown length
* 19519 20710: contig of 1192 bp in length
* gap of unknown length
* 20711 21932: contig of 1222 bp in length
* gap of unknown length
* 21933 23606: contig of 1674 bp in length
* gap of unknown length
* 23607 25328: contig of 1722 bp in length
* gap of unknown length
* 25329 26649: contig of 1321 bp in length
* gap of unknown length
* 26650 27849: contig of 1200 bp in length
* gap of unknown length
* 27850 29135: contig of 1286 bp in length
* gap of unknown length
* 29136 30700: contig of 1565 bp in length
* gap of unknown length

* 30701 31877: contig of 1177 bp in length
* gap of unknown length
* 31878 33065: contig of 1188 bp in length
* gap of unknown length
* 33066 34267: contig of 1202 bp in length
* gap of unknown length
* 34268 35474: contig of 1207 bp in length
* gap of unknown length
* 35475 36650: contig of 1176 bp in length
* gap of unknown length
* 36651 37859: contig of 1209 bp in length
* gap of unknown length
* 37860 39072: contig of 1213 bp in length
* gap of unknown length
* 39073 40406: contig of 1334 bp in length
* gap of unknown length
* 40407 41630: contig of 1224 bp in length
* gap of unknown length
* 41631 42864: contig of 1234 bp in length
* gap of unknown length
* 42865 44200: contig of 1336 bp in length
* gap of unknown length
* 44201 46087: contig of 1887 bp in length
* gap of unknown length
* 46088 48227: contig of 2140 bp in length
* gap of unknown length
* 48228 49518: contig of 1291 bp in length
* gap of unknown length
* 49519 50910: contig of 1392 bp in length
* gap of unknown length
* 50911 52420: contig of 1510 bp in length
* gap of unknown length
* 52421 53613: contig of 1193 bp in length
* gap of unknown length
* 53614 55004: contig of 1391 bp in length
* gap of unknown length
* 55005 56300: contig of 1296 bp in length
* gap of unknown length
* 56301 57505: contig of 1205 bp in length
* gap of unknown length
* 57506 58686: contig of 1181 bp in length
* gap of unknown length
* 58687 60407: contig of 1721 bp in length
* gap of unknown length
* 60408 61632: contig of 1225 bp in length
* gap of unknown length
* 61633 62880: contig of 1248 bp in length
* gap of unknown length
* 62881 64220: contig of 1340 bp in length
* gap of unknown length
* 64221 65420: contig of 1200 bp in length
* gap of unknown length
* 65421 66605: contig of 1185 bp in length
* gap of unknown length
* 66606 68682: contig of 2077 bp in length
* gap of unknown length
* 68683 70712: contig of 2030 bp in length
* gap of unknown length
* 70713 71895: contig of 1183 bp in length
* gap of unknown length
* 71896 73462: contig of 1567 bp in length
* gap of unknown length
* 73463 74806: contig of 1344 bp in length
* gap of unknown length
* 74807 76205: contig of 1399 bp in length
* gap of unknown length
* 76206 77580: contig of 1375 bp in length
* gap of unknown length
* 77581 79081: contig of 1501 bp in length
* gap of unknown length
* 79082 80314: contig of 1233 bp in length
* gap of unknown length
* 80315 81899: contig of 1585 bp in length

* 81900 83140: gap of unknown length
* contig of 1241 bp in length
* 83141 84665: gap of unknown length
* contig of 1525 bp in length
* 84666 85888: gap of unknown length
* contig of 1223 bp in length
* 85889 87155: gap of unknown length
* contig of 1267 bp in length
* 87156 88435: gap of unknown length
* contig of 1280 bp in length
* 88436 89955: gap of unknown length
* contig of 1520 bp in length
* 89956 91277: gap of unknown length
* contig of 1322 bp in length
* 91278 92513: gap of unknown length
* contig of 1236 bp in length
* 92514 93917: gap of unknown length
* contig of 1404 bp in length
* 93918 95448: gap of unknown length
* contig of 1531 bp in length
* 95449 96800: gap of unknown length
* contig of 1352 bp in length
* 96801 97978: gap of unknown length
* contig of 1178 bp in length
* 97979 99179: gap of unknown length
* contig of 1201 bp in length
* 99180 100437: gap of unknown length
* contig of 1258 bp in length
* 100438 101827: gap of unknown length
* contig of 1390 bp in length
* 101828 103117: gap of unknown length
* contig of 1290 bp in length
* 103118 104550: gap of unknown length
* contig of 1433 bp in length
* 104551 105854: gap of unknown length
* contig of 1304 bp in length
* 105855 107523: gap of unknown length
* contig of 1669 bp in length
* 107524 109208: gap of unknown length
* contig of 1685 bp in length
* 109209 110430: gap of unknown length
* contig of 1222 bp in length
* 110431 111662: gap of unknown length
* contig of 1232 bp in length

Query Match 72.6%; Score 675; DB 2; Length 212936;
Best Local Similarity 98.1%; Pred. No. 4.4e-164; Indels 17; Gaps 17;
Matches 862; Conservative 0; Mismatches 0

Qy 52 TCTTTTACCGCCCGCCGACAGCTACAGATGAGCGTTTCGTCCATCAGTACAGCGTA 111
Db 168976 TCTTTTACCGCCCGCCGACAGCTACAGATGAGCGTTTCGTCCATCAGTACAGCGTA 169034

Qy 112 TCAAACTGGAAGATGAGTTGCGAGTAAAGCTGTTAAACGTTAGCAGCAGCAGCATTGGC 171
Db 169035 TCAAACTGGAAGATGAGTTGCGAGTAAAGCTGTTAAACGTTAGCAGCAGCAGCATTGGC 169093

Qy 172 CTGACCGAAGCCGCTAGAAATTTTACTACAGGGTGCCTGCTATGCTTTCATGAAGTGCAG 231
Db 169094 CTGACCGAAGCCGCTAGAAATTTTACTACCA-GGCTGCGTGTGATGCTTTCATGAAGTGCAG 169152

Qy 232 GATGTTTCATGAGCACTGTATGCTTCAATACACCCCATCGGAGCGTACGATTGGC 291
Db 169153 GATGTTTCATGAGCACTGT-TGCCTTCAATAACACCCCATCGGAGCGTACGATTGGC 169211

Qy 292 TGTTCCTTCAACTATGGCAGAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAGAA 351
Db 169212 TGTTCCTC-ACATGAGCAAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAG-A 169269

Qy 352 TACCCAGGTTGAGCGTCAATCTGGTTACCGAAATTCAGCCCGCCGACCTGATTGCCGAC 411
Db 169270 TACCCAGGTTGAGCGTCAATCTGGTTACCGAAATTCAGCCCGCCGACCTGATTGCCGAC 169328

QY	412	GGTCTGGATGGTGGATCGCGGTGCGCGGTGAGGATTCAGCCTGTTTCCGCGCGT	471
Db	169329	GGTCTGGATGGTGGATCGCGGTGCGCGGTGAGGATTCAGCCTGTTTCCGCGCGT	169387
QY	472	CTGGCGCGGATCCCAATGGTGGTGGCGCGGAAAGCTATCTCACAAATACGGCATA	531
Db	169388	CTGGCGCGGATCCCAATGGTGGTGGCGCGGAAAGCTATCTCACAAATACGGCATA	169446
QY	532	CCG3AAAACCCCGCGATTTGAGTAGTATTCATGGCTTGAATACAGCGTCGCGCCGAC	591
Db	169447	CCG3AAAACCCCGCGATTTGAGTAGTATTCATGGCTTGAATACAGCGTCGCGCCGAC	169505
QY	592	AATGAATTTGAATGATCGCACCGAAGGATCTCGACTCGCTGATCCCAAGGAAGA	651
Db	169506	AATGAATTT-ACATGATCGCACCGAAGGATCTCGACTCGCTGATCCCAAGGAAG-	169563
QY	652	TTTGTGACTAATGATCCGATGACGCTGGTGGCTGGCTGACGCGGGTGGCGGATCGCC	711
Db	169564	TTTGTGACTAATGATCCGATGACGCTGGTGGCTGGCTGACGCGGGTGGCGGATCGCC	169622
QY	712	TACGTGCGCTGATGGTGGTGAACAGAGATCAATCGTGGGAGCTGAGATCCTGCTG	771
Db	169623	TACGTGCGCTGATGGTGGTGAACAGAGATCAATCGTGGGAGCTGAGATCCTGCTG	169681
QY	772	CGCGGTACAGTACGATCCAGCCCGGTTTATGCGTTATATACCGAAAAGATAAGCTG	831
Db	169682	CGCGGTACAGTACGATCCAGCCCGG-TTATGCGTTATATACCGAAAAGATAAGCTG	169740
QY	832	CGCTGAAGTACAGTCTGATCACTCGCTGACGCGGATTTTGTGAGTGGGTAAA	891
Db	169741	CGCTGAAGTACAGTCTGATCACTCGCTGACGCGGATTTTGTGAGTGGGTAAA	169799
QY	892	TTGTTTCAGAGATGACGCGGCGCGGAAAGAGATAA	930
Db	169800	TTGTTTCA-GAGATGACGCGCGCGGAAAGAGATAA	169837
RESULT 9			
LOCUS	AL627278	258050 bp	DNA linear BCT 04-JUL-2003
DEFINITION	Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18, complete chromosome, segment 14/20.		
ACCESSION	AL627278	AL513382	
VERSION	AL627278.1	GI:16504263	
KEYWORDS	Salmonella enterica subsp. enterica serovar Typhi		
SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.		
ORGANISM	1 (bases 1 to 258050)		
REFERENCE	Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G., Sebahia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connor, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S. and Barrell, B.G. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18 Nature 413 (6858), 848-852 (2001)		
JOURNAL	21534947		
MEDLINE	11677608		
PUBMED	2 (bases 1 to 258050)		
REFERENCE	Parkhill, J.		
AUTHORS	Direct Submission		
TITLE	Submitted (25-OCT-2001) Submitted on behalf of the Salmonella sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK		
JOURNAL	B-mail: parkhill@sanger.ac.uk		
COMMENT	Notes: Details of S. typhi sequencing at the Sanger Centre are available on the World Wide Web.		

FEATURES		(URL, http://www.sanger.ac.uk/Projects/s_typhi/).
source	Location/Qualifiers	
	1..258050	/organism="Salmonella enterica subsp. enterica serovar Typhi"
		/mol_type="genomic DNA"
	/strain="CT18"	/db_xref="taxon:90370"
gene	69..1544	/gene="tolC"
		/note="synonym: STY33364"
CDS	69..1544	/gene="tolC"
		/note="Similar to Salmonella enteritidis outer membrane protein TolC precursor tolC SW:TOLC_SALEN (Q54001) (491 aa) fasta scores: E(): 0.99, 0% id in 491 aa, and to Escherichia coli outer membrane protein tolC precursor tolC or mcb or muka or refi SW:TOLC_ECOLI (P02930) (495 aa) fasta scores: E(): 0.89, 7% id in 495 aa Orthologue of E. coli tolC (TOLC_ECOLI); Fasta hit to TOLC_ECOLI (495 aa), 90% identity in 495 aa overlap"
		/codon_start=1
		/transl_table=11
		/product="outer membrane protein TolC precursor"
		/protein_id="CAD07712.1"
		/db_xref="GI:16504264"
		/db_xref="GOA:Q8Z3N8"
		/db_xref="TrEMBL:Q8Z3N8"
		/translation="MQMKLLPILIGLSLGSFSLSQENLMQVYQOARLNPFLRKSADRAAPFKINEARSPILLPOLGADVTYSNGYRDANGINETSASLQTLQFLDM SKWRGLTQEKAGIQDVTYQDQTLILNTANAYKVLNAIDVLSYTAQAKYLYRO LQQTQRFNVVAITDVONARAQDVTLANEVTARNLDNAVEIRLQVTVGYDELA SLNVEHFKDKPAVNALKEAENRNLISLQARLSQDLAREIQIRQDGHLPMLTA STGIDTSYSGKTSNAQYDDSMQNGKINGLNLSPLFYGGMVNSQVQKQAFNVGAS EQLSAHRSVQVTVSSFNINASISINAYKQAVVSACSLDAVEAGYSVGTRTIVD VLDATTTLYDAKQALNARYTILNQLNIKYALGTLEQDLALANSLTGKPTSPES VLPATPEQDAADGYNAHSAAPAVQPTAARANSNNGNPPRH"
gene	1757..2428	/gene="STY3365"
CDS	1757..2428	/gene="STY3365"
		/note="Orthologue of E. coli ygiB (YGIH_ECOLI); Fasta hit to YGIH_ECOLI (234 aa), 97% identity in 223 aa overlap. Contains a possible N-terminal signal sequence."
		/codon_start=1
		/transl_table=11
		/product="possible lipoprotein"
		/protein_id="CAD07713.1"
		/db_xref="GI:16504265"
		/db_xref="TrEMBL:Q8XG21"
		/translation="VKRTKSIHHSFRKSWSAHLPVALAVTAVFVLACGCKSDSTV SLVQNAADCSAANPGKASCTTAYNNALKEAPRTAPKATREDCAEFCGCGQOAPA QGMAEPNQAQACGSGSFWFMELMAGYMWGRMGCGGAGFAQQFPLSSKNPSPAYGKY TDAAGKNYGAAPGRTTVPKTMAPKPAITITVINGGFGESVAKQSTVQRSAAAGTST RSMGG"
misc_feature	1835..1867	/gene="STY3365"
		/note="PS00013 Prokaryotic membrane lipoprotein lipid attachment site"
gene	2434..3397	/gene="STY3366"
CDS	2434..3397	/gene="STY3366"
		/note="Fasta hit to YJFC_ECOLI (387 aa), 50% identity in 392 aa overlap Orthologue of E. coli ygiC (YGIH_ECOLI); Fasta hit to YGIH_ECOLI (386 aa), 94% identity in 386 aa overlap"
		/codon_start=1
		/transl_table=11
		/product="conserved hypothetical protein"
		/protein_id="CAD07714.1"
		/db_xref="GI:16504266"
		/db_xref="TrEMBL:Q8XFP6"

```
gene
/translation="VERVSIITRPPWRDKATEYGFNFTMYGEPYWCEDAYKJTLAQ
VEKLSVTAELQMCILKVVVERVIADELMETKFRPKHTWGFVRSQWQQLSLYRLD
LAWDGIKPKLENNADPTSLYEAATQWLEQINAGNLPESQDFNLSQKLEIE
RFAERQYFOLLHLTCRDTVEDRTQIQDCAAEAIATEFLYIDDLGLGKQ
FDLQOQIANFLKLYPWFMLREMFSTKLEDAVGRLEPAWKSIISNKLPLIWM
PDHNLPLPAYAEDEHPMDKVVKPIPSREGANVSIENGKTIIESVEGYPGEGMI
VQOYPLPKGDSVTLIGSLWINDQPAIGIREDRALITQLDSRPHIFVEG"
complement(3659..4489)
/gene="STY3367"
CDS
/gene="STY3367"
complement(3659..4489)
/gene="STY3367"
/notes="Orthologue of E. coli ygid (YGID_ECOLI); Fasta hit
to YGID_ECOLI (271 aa), 87% identity in 266 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD0715.1"
/db_xref="GI:16504267"
/db_xref="GOA:Q8Z3N7"
/db_xref="TrEMBL:O8Z3N7"
/translation="MVMLTKLSCKDNIMSLTCLPFLGHGSPMVLDDNDYTRWR
RIGELAPRQALIVVSALWYINGTGTWATERPQTLHDFGGFQALYDHYPAQSPAL
ARLVELLAPVVALDKAWGDHGSWGLIKMYPNADIPMVQSVDSVTKPAWHFEM
GRLKATLDEGYMLVASGNVHLRTVRHGDNIYPYWAASFDFVKANLTWQGFVEQ
HPLVNYLQHEGGLSNPTPEHFLPLLYVLGAWDGKEPITIPVDGIEMSGISMLSVQVG"
4588..5361
/gene="STY3368"
CDS
/notes="STY3368"
/notes="Orthologue of E. coli ygie (YGI_ECOLI); Fasta hit
to YGI_ECOLI (257 aa), 93% identity in 257 aa overlap.
Contains multiple possible membrane spanning hydrophobic
domains"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD0716.1"
/db_xref="GI:16504268"
/db_xref="GOA:Q8XGR4"
/db_xref="Swiss-Prot:Q8XGR4"
/translation="MSVPLLIITLLAGATFTGFLGVLPQKPSNRVLAFSLGPAAGIM
LILSLEMPLPAALDTGEMSPVLGYGNFIIGLGYGLDLRLLPHAPQDLVQKQOPLP
GSIXTALLITGLSHNPEEGIATFTVASNLELFGIALVALHNLPEGLVAGVP
YAATGSKRTAIFWAGISQWAEILGGLVLAHLIIGLSVSPVMAAIAVAAGINVALSVD
ELMPLAKSIDPNNPSYGVLCGMSIMGLSLVILQIGIG"
6004..7800
/gene="STY3370"
CDS
/notes="STY3370"
/notes="Synonym: asst"
6004..7800
/gene="STY3370"
/notes="Similar to Klebsiella pneumoniae arylsulfate
sulfotransferase asst TR:P97036 (EMBL:U32616) (598 aa)
fasta scores: E(): 0, 86.8% id in 598 aa"
/codon_start=1
/transl_table=11
/product="probable arylsulfate sulfotransferase"
/protein_id="CAD0717.1"
/db_xref="GI:16504269"
/db_xref="GOA:Q8Z3N6"
/db_xref="TrEMBL:O8Z3N6"
/translation="WEDQYRKTILAGAVALTCLGTLAATPAAGFCPAQPAQKLGAVV
DYGNAFTVLELHSHIISDVVKVTHGKRGKVPVTVIGKRESLETYDGIPIGLYQ
KFAVNPTVEKNGAMKDDYVQTSIAVNHMDNRSISDQTKVKIVKAPGFDRLY
LYNTHFTPQGAEPFWHGEKKNAGILDAPGALPFDIAPIYTFVVDTOGEYRWLD
QDTFYNGHDMINKRGYLMGIRTPRGFTTAVQGOHWYEFDMGQILADHKLPRGFLD
ASHESIETVNGVLLVRGKDYRKEDGTHVHTIRDOIIEVDKSGRVDVWDLTKILD
MEDALLDAGACVNDLAHAGQAKLEPDTPYGDALGVAGRNWAVNSIADK
DSDILSRHQIVIGIKGDKQVWILAPSKGNKQLASKLLKXPVDDHGKPLTCBNGK
KQDQDFDTYTHQTAWSKGLTVFDNGDGRLEQPALPNTKSRFVEYKIDKRGY
VQOWVEYKGRGKDYFSITSVVEYKQDRDNTFPGGGINLFDVGKPTVGLKNEIDYK
TKEVKBIDVLSDKRNQTHRYALLVHPITQMFK"
7820..8491
gene
```

```

/gene="dsba"
/notes="synonym: STY3371"
7820..8491
CDS
/notes="dsba"
/notes="Similar to Enterobacter amnigenus disulfide
isomerase dsba TR:09XDPI (EMBL:AF012826) (222 aa) fasta
scores: E(): 0, 90.1% id in 222 aa"
/codon_start=1
/transl_table=11
/product="disulfide isomerase"
/protein_id="CAD0718.1"
/db_xref="GI:16504270"
/db_xref="GOA:O8Z3N5"
/db_xref="TrEMBL:O8Z3N5"
/translation="MSKKNITSLFKSVLITLALVTPFAASAFTEGTDYMWLEKPIPA
DKTLIKVFSIACFPCKYDKAVTGPVSDKVDLVTFTPEHLTKGYSQASVFAVL
IAKDKAAGISLFDKASKQFKKAFWATYHDKERWSDGKDPAAFIKTGLDAAGMSQA
DFEALAKDPAVQSTLEKRWAAAYDAKIQGVPAVVVNGKYLITKNIKSIDSMALVRE
LATKK"
misc_feature
7847..8488
Query Match 72.1%; Score 670.8; DB 1; Length 258050;
Best Local Similarity 82.6%; Pred. No. 5.5e-163;
Matches 769; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
QY 1 ATGGAACGACATAAAACGATGTCGGTGTTCGCCAAGTAGTTCGAATTTGGCTCTTTTACC 60
Db 180896 ATGGAACGATATAAACGAATGTCGGTATTCGCCAAGTTGTTGAGTTTGGCTCTTTTACC 180955
QY 61 GCCGCCCCAGACGAGCTACAGATGAGCGTTTTCCTCATCATAGTCAGACGGTATCAAAACTG 120
Db 180956 GCCGAGCCAGACGAGCTGCAATGAGCGTATCGCTCCATCAGCAGACTGTGGGAACTG 181015
QY 121 GAAGATGAGTTGCGAGGTAAAGCTGTTAAACCGTAGCACGCGAGCATTTGCCCTGACCGAA 180
Db 181016 GAAGATGAGTTCAGGTAAATTTGCTCACCGCAGCACGCGAGCATTTGGGCTCACCGAA 181075
QY 181 GCCGTAAGATTTACTACAGGCGTCGCGTGTATGCTTCATGAAGTGCAGGATGTTTCAT 240
Db 181076 GCTGGAATAATCTATTATCAGGCGTGTGCTGCTATGCTGATGAAGTGCAGGATGTTTCAC 181135
QY 241 GAGCACTGTATGCTTTCATATACACCCCATCGGACGCTACGCGTATGCTGCTCTTCTTCA 300
Db 181136 GAGCAGCTTTACGCTTTTAAATACAGCGCTATCGGTACATTCGCTATCGGCTGCTTCTTCA 181195
QY 301 ACTATGSCACAAAATGTTCTCGCGGCGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360
Db 181196 ACCATGCGCAAAATGCTGCGCGGCTTACCGCAAAATATTATAAGAGTATCCCGGT 181255
QY 361 TTGAGCGTCAATCTGTTTACCGAAATTCAGCCCCCGACCTGATTCGCCAGCGTCTGGAT 420
Db 181256 CTGCGGTCATCTGTTGACAGGAATTCGCGCCCCGGAATTAATTCGCCAGCGGCTGGAC 181315
QY 421 GTGTGATCCGCGTCGCGCGGTTGCAAGGATTCAGAGCTGTTTCCCGCGCTCTGGCGCG 480
Db 181316 GTGTGATCCGCTGGCGCGCTGCAAGACTCCAGCTGTTTCCCGTGCAGTCTGGCGCG 181375
QY 481 ATGCCAATGTTGTTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
Db 181376 ATGCCGATGTTGTTGTCGCGCGGCAAGCCCTACCTCGCTCAGTACGCGGCGGAAAAA 181435
QY 541 CCGCGCGATTTGAGTAGTCTATTCATGGCTGATACAGCGTGGCGCGCGACRATGATTT 600
Db 181436 CCGCGCGATCTCAGTAGCCATTCCTGCGTGAATATAGCGTGGCGCGCGATATAGTTTT 181495
QY 601 GAATGATCCGACCGGAAGGATCTCGACTCGGCTGATCCCAAGAAAGATTTGTGACT 660
Db 181496 GAATTTATGGCGCGGAAGAAATTTCTACCGACTGATTTCCACAGGGCGGTTTCGCTAC 181555
QY 661 AATGATCCGATGACGCTGTTGCGCTGAGCGGGGTGCCGGGATCGCTACGTCGCGG 720
Db 181556 AACGATCCGATGACGCTGTTGCTGCTGACGCGCGGAAACCGGAAATCGCTTATGTGCGG 181615
```

```
QY 721 CTGATGTCGGTGCATCAACGAGATCAATCGTGGGAGCTGAGATCCTGTCGCCGCTTAC 780
Db 181616 CTGATGTCGGTGCATGATGATGATCAATCGTGGGATCGGATCCTGCTGCCTAT 181675
QY 781 CAGTCAGATCCAGCCCGGTTTATGCGTTATATACCGAAAGATAGTCGCCGCTGAAG 840
Db 181676 CAGTCAGATCCCGGTCGCGTGTACCGGCTGTACCGGAAAGATAGTCGCCGCTTAA 181735
QY 841 GTACAGCTGTCGATCAACTCGCTGACGCGATTATTGTTGAGCTCGGTAAATGTTTCAG 900
Db 181736 GTGACAGCTGTCGATCACTGATTAACCGCTTAATGTTGAGCTGTCGACACCTGTTTCAG 181795
QY 901 GAGATCAGCGGCGCGGGAAGAGAGTAA 930
Db 181796 GGGATCAGCGGCGGAGGGAAGAAATAA 181825

RESULT 10
AE016845
LOCUS
DEFINITION
  299991 bp DNA linear BCT 21-MAR-2003
  Salmonella enterica subsp. enterica serovar Typhi Ty2, section 12
  of 16 of the complete genome.
ACCESSION
  AE016845 AE014613
VERSION
  AE016845.1 GI:29139182
KEYWORDS
  Salmonella enterica subsp. enterica serovar Typhi Ty2
  Salmonella enterica subsp. enterica serovar Typhi Ty2
  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Salmonella.
SOURCE
  1 (bases 1 to 299991)
  Deng, W., Liou, S.-R., Plunkett III, G., Mayhew, G. F., Rose, D. J.,
  Burland, V., Kodoyianni, V., Schwartz, D. C. and Blattner, F. R.,
  Comparative Genomics of Salmonella enterica serovar Typhi Strains
  Ty2 and CT18
  J. Bacteriol. 185 (7), 2330-2337 (2003)
  22531367
  12644504
REFERENCE
  1 (bases 1 to 299991)
  Deng, W., Liou, S.-R., Plunkett, G. III, Mayhew, G. F., Rose, D. J.,
  Burland, V., Kodoyianni, V., Schwartz, D. C. and Blattner, F. R.,
  Direct Submission
  Submitted (25-SEP-2002) Laboratory of Genetics, University of
  Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
  Location/Qualifiers
  1..299991
  /organism="Salmonella enterica subsp. enterica serovar
  typhi Ty2"
  /mol_type="genomic DNA"
  /strain="Ty2"
  /serovar="Typhi"
  /sub_species="enterica"
  /db_xref="taxon:209261"
  complement(83..709)
  /gene="ftsJ"
  /locus_tag="t3214"
  complement(83..709)
  /gene="ftsJ"
  /locus_tag="t3214"
  /note="corresponds to STY3475 from Accession AL513382:
  Salmonella typhi CT18"
  /codon_start=1
  /transl_table=11
  /product="cell division protein"
  /protein_id="AAC07050.1"
  /db_xref="GI:29139183"
  /translation="MTGKKRSASSENLOEHFSKYYQAOAKGRLSPAFKLEIQO
  SKLKPKPMVTVDLGAARFGSQYVVTQIGKGRLLIACDLPMDFIVGVDFLQDFRD
  ELVMKALLERVDKSVQVWMSDAPNMGSTPAVDIPRAVYLVEALDMCDVLAAPGGS
  FVVKVFGEGDFEIRELSFTKVKVRKPDSSRRARSREVIVATGR"
  797..1129
  /locus_tag="t3215"
  797..1129
  /locus_tag="t3215"
  /locus_tag="t3215"

gene
  source
  /organism="Salmonella typhi CT18"
  /product="probable GTP-binding protein"
  /protein_id="AAC07054.1"
  /db_xref="GI:29139187"
  /translation="MKFVDEASILVAVGGNGCVSFRREKVIKPGGPDGGDGGGG
  VWWEADENLTLDYRFEKSPRAESGONGASDCTGKQKDVTKVPVGRTRVIDOGTG
  ETNGDMTRKGRQLLVAKGWHGLGNTRFKSSVNRTPRQNTNTPGKRDRLLEMLLA
  DVMLGMPNAGKSTFIRVSAAPKPVADYPTFTLVPSLGVVRMSKSFVADIPGLI
  EGAAEGAGLGIIRFKHLERLVLLDIDIDPSDPVENARIIIGELEKYSQDLAAK
  PRLVFNKIDLMKSEAEKAKAIAEALGWEGKYLISAASQLGKDLCDVMYTFIIE
  NPTAQAEAKQKPEVFEMDDVHRQQLAEVEEDADDWDDEDEGEVEFYIKR"
  complement(4761..5726)
  /locus_tag="t3219"
  complement(4761..5726)
  /locus_tag="t3219"

CDS
  /note="corresponds to STY3475 from Accession AL513382:
  Salmonella typhi CT18"
  /codon_start=1
  /transl_table=11
```

```
/notes="corresponds to STY3476 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO70751.1"
/db_xref="GI:29139184"
/translation="MTRFOSQRQKYTNLSTKQKHLKGLAHPKVPVVLGNNGLTE
GVLAETEOALEHHELIKVIASEDRETKTLIVDAIVRETGACNVQVIGKTLVLRPTK
ERKISLPR"
complement(1285..1761)
/gene="greA"
/locus_tag="t3216"
complement(1285..1761)
/gene="greA"
/locus_tag="t3216"
/notes="corresponds to STY3477 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="transcription elongation factor"
/protein_id="AAO70752.1"
/db_xref="GI:29139185"
/translation="MQAIPMTLRGAEKLEELDFKSVRRPPIIAIAEAREHGDLKE
NAEYHAARQQGFCGRIKDIKAKLSNAQVIVDTKMPNNGRVIFGNTVTLNLDTEE
QYTRIVGDDDEADFQKNLISVNSFIARGLIGKEQDDVVVKTGGDVEVEVLKVEYL"
2009..3442
/locus_tag="t3217"
2009..3442
/locus_tag="t3217"
/notes="corresponds to STY3479 from Accession AL513382:
Salmonella typhi CT18; D-alanyl-D-alanine
carboxypeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AAO70753.1"
/db_xref="GI:29139186"
/translation="MRFSEFIILGTSTAFSVQAAINIDEVVKOLPAGANLALMVOKIG
APAPDIHSQQMALPASQKVITALLALILQLQDPFRFTTLETGKQVNDGILKGDLI
ARFGDPTLRQDIRNMVATLKKSGVTQIDGNVLDTISFASHDKAPGWNDDLTQCF
SAPPAAIYDRNCFVSLSYSAQKPNDAFIRVASYPTVMFSQVTLPRGSADAYCE
LDVPGDLNRYLTGCLPQADPLPAFQIDQASVAGAILKQELKEAGITVYRGLLR
QTVNBPPTIVASKQAPLHLLKIMLKKSDNMIADTVPRMIGHVRFNVPFWRAGSD
AVRQILRQOAGIDIGNTIADGSGLSRHNLIAPATMMQVLOVIAOHDNENLFIIMPL
AGYDGSLOYRAGLHQGVGVKYSAGTSQGVYNLAGFITTASGQEMAFVQVLSGVAV
PAPQNRRIPLVRFESRUYKDIYQNN"
complement(3573..4745)
/locus_tag="t3218"
complement(3573..4745)
/locus_tag="t3218"
/notes="corresponds to STY3480 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="probable GTP-binding protein"
/protein_id="AAO70754.1"
/db_xref="GI:29139187"
/translation="MKFVDEASILVAVGGNGCVSFRREKVIKPGGPDGGDGGGG
VWWEADENLTLDYRFEKSPRAESGONGASDCTGKQKDVTKVPVGRTRVIDOGTG
ETNGDMTRKGRQLLVAKGWHGLGNTRFKSSVNRTPRQNTNTPGKRDRLLEMLLA
DVMLGMPNAGKSTFIRVSAAPKPVADYPTFTLVPSLGVVRMSKSFVADIPGLI
EGAAEGAGLGIIRFKHLERLVLLDIDIDPSDPVENARIIIGELEKYSQDLAAK
PRLVFNKIDLMKSEAEKAKAIAEALGWEGKYLISAASQLGKDLCDVMYTFIIE
NPTAQAEAKQKPEVFEMDDVHRQQLAEVEEDADDWDDEDEGEVEFYIKR"
complement(4761..5726)
/locus_tag="t3219"
complement(4761..5726)
/locus_tag="t3219"

gene
  source
  /organism="Salmonella typhi CT18"
  /product="probable GTP-binding protein"
  /protein_id="AAC07054.1"
  /db_xref="GI:29139187"
  /translation="MKFVDEASILVAVGGNGCVSFRREKVIKPGGPDGGDGGGG
  VWWEADENLTLDYRFEKSPRAESGONGASDCTGKQKDVTKVPVGRTRVIDOGTG
  ETNGDMTRKGRQLLVAKGWHGLGNTRFKSSVNRTPRQNTNTPGKRDRLLEMLLA
  DVMLGMPNAGKSTFIRVSAAPKPVADYPTFTLVPSLGVVRMSKSFVADIPGLI
  EGAAEGAGLGIIRFKHLERLVLLDIDIDPSDPVENARIIIGELEKYSQDLAAK
  PRLVFNKIDLMKSEAEKAKAIAEALGWEGKYLISAASQLGKDLCDVMYTFIIE
  NPTAQAEAKQKPEVFEMDDVHRQQLAEVEEDADDWDDEDEGEVEFYIKR"
  complement(4761..5726)
  /locus_tag="t3219"
  complement(4761..5726)
  /locus_tag="t3219"

CDS
  /note="corresponds to STY3481 from Accession AL513382:
  Salmonella typhi CT18"
  /codon_start=1
  /transl_table=11
```

/product="putative membrane protein"
/protein_id="AAO70755.1"
/db_xref="GI:29139188"
/translation="MKQAGGICILLATTCMCWGPALPIAMQVLEWEPSTIVYRFL
MASIGLGAIVAVKRLPLPIKPRKRLVLLAIATCGLFGNFIPLFSSLSQYLSPASQ
VIGQSPVGMVAVSFIKIKRGRQV:GALMLGSLVYFFNTSLIEFTLRLTYWG
VIFGGAAMVSYGAKVLRRLASQIIFLLYTLCTIALLLPLAKPMVIAQLSDWQ
LACILFCGLNTLVGCAAEARWAAQVSAIITLPLTLLFDLILSMAMPDFAR
PVLNLLGLGARFVVAGAMYSAGHRIWGGLRKHETVVSQPRSGE"
complement(5856..6113)
/gene="rpnA"
/locus_tag="t3220"
complement(5856..6113)
/gene="rpnA"
/locus_tag="t3220"
/note="corresponds to STY3482 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L27"
/protein_id="AAO70756.1"
/db_xref="GI:29139189"
/translation="WAHKKAGSTENGSDSEAKRLGVKRGGEAVLAGSIIVRQGTG
FIAGNVGGRDHTLPAKADGKVFVKPKRKYISIAE"
complement(6133..6444)
/gene="rplU"
/locus_tag="t3221"
complement(6133..6444)
/gene="rplU"
/locus_tag="t3221"
/note="corresponds to STY3483 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="50S ribosomal subunit protein L21"
/protein_id="AAO70757.1"
/db_xref="GI:29139190"
/translation="WFAVPSGKGKRVSEGTVLEKLDIATGTIFPAEVLMIANG
EEVKGVPVDPDGGVIAEVAHGRKEKIVFRKRRKHYRQKQGRHQPFTDKITGIS
A"
6703..7674
/gene="ispB"
/locus_tag="t3222"
6703..7674
/gene="ispB"
/locus_tag="t3222"
/note="corresponds to STY3484 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="octaprenyl-diphosphate synthase"
/protein_id="AAO70758.1"
/db_xref="GI:29139191"
/translation="WNLEKINELTAQDMAGVNATILEQINDSDVQLINOLGYIISGGG
KRIRPMIAVLAARAVYQGNHVTIAALI:EFTHATLLDHDVDESDMRGKATNAA
FGNAASVLGDFITRAFOFMT:SLGSLKVLVSEAVNVIAEGLVQLMNVNDPDITE
ENMYRVIYKATLPEAAQCSGILAGCTPECKGLQDXGYLGTAFOLIDLDLYSA
DEHLKKNVDLNGKPTLPULLHMRHGTPEQSMIRTAIEQNGRHLLEPVLBMT
TCGSUWTKQRAEEBKAISALQILPDTFWREALIGLHAIAVQDR"
7307..8194
/gene="nlp"
/locus_tag="t3223"
7307..8194
/gene="nlp"
/locus_tag="t3223"
/note="corresponds to STY3485 from Accession AL513382:
Salmonella typhi CT18"
/codon_start=1
/transl_table=11
/product="Ner-like regulatory protein"
/protein_id="AAO70759.1"
/db_xref="GI:29139192"
/translation="MESKLIIDWHPADIAGLRKKGTSMASRRNGLSSSTIANALTR

Query Match	72.1%;	Score 670.8;	DB 1;	Length 299991;
Best Local Similarity	82.6%;	Pred. No. 5.6e-163;		
Matches	768;	Conservative	0;	Mismatches 162; Indels 0; Gaps 0;
Qy	1	ATGGAACGATATAAACCCATGTCGGTGTTCGCAAAAGTAGTGAATTTGGCTCTTTTACC	60	
Db	65727	ATGGAACGATATAAACCCATGTCGGTGTTCGCAAAAGTAGTGAATTTGGCTCTTTTACC	65786	
Qy	61	GCGCCGCCACAGAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAAACTG	120	
Db	65787	GCGCCGCCACAGAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAAACTG	65846	
Qy	121	GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCACACGCGAGCTTGGCCCTGACCGAA	180	
Db	65847	GAAGATGAGTTGCGAGTAAAGCTGTTAAACCGTAGCACACGCGAGCTTGGCCCTGACCGAA	65906	
Qy	181	GCCGGTAGAATTTACTACGAGGCTGCCGTGCTATGCTTCATGAAGTCAGAGTAGTTCAT	240	
Db	65907	GCTGGAATAATCTAATATCAGGCGTGTGTCGTATGCTGATGAAGTCAGAGTAGTTCAC	65966	
Qy	241	GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTAGCGATTGGCTGTCTTCA	300	
Db	65967	GAGCAGCTTTACGCTTTTAAATACACGCTATCGGTACATGCGTATCGGCTGTCTTCA	66026	
Qy	301	ACTATGGCACAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAAAGTAATCCCAAGT	360	
Db	66027	ACCATGGCGAAAATGTTCTCGCGGGCTTACCGCAAAAATATTATAAAGAGTAGTTCGCGT	66086	
Qy	361	TTGAGCGTCAATCTGTTTACCGGAAATTCAGACGCCCGACCTGATTCGCGACGGTCTGGAT	420	
Db	66087	CTGCGGTCATCTGTTGACAGAAATTCGCGCCCGGATTTAATTCGCGACGGGCTGGAC	66146	
Qy	421	GTGTGATCGCGCTCGCGCGCTGTCAGGATTCAGAGCTGTTTCCGCCCTCTCTGGCGCG	480	
Db	66147	GTGTGATCGCGCTCGCGCGCTGTCAGAGCTTCCAGCGCTGTTTCCGCCCTCTCTGGCGCG	66206	
Qy	481	ATGCCAATGTTGTTGCGCGCGGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAA	540	
Db	66207	ATGCCAATGTTGTTGCGCGCGGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAA	66266	
Qy	541	CCCCCGGATTTGAGTAGTCATTTATGGCTTGAATACAGCGTGGCGGCCGACAAATGAATTT	600	
Db	66267	CCCCCGGATTTGAGTAGTCATTTATGGCTTGAATACAGCGTGGCGGCCGACAAATGAATTT	66326	
Qy	601	GAACTGATCGCACCGGAAGGATCTCGACTCGGCTGATCCACAAAGGAATTTGTGACT	660	
Db	66327	GAACTGATCGCACCGGAAGGATTTCTACCGGACTGATCCACAGGGGCGGTTTGTGACT	66386	
Qy	661	AATGATCCGATGACGCTGCTGCTGCTGACGCGGGTGGCGGATCGCTACGTCGG	720	
Db	66387	AACGATCCGATGACGCTGCTGCTGCTGACGCGGGTGGCGGATCGCTACGTCGG	66446	
Qy	721	CTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATTCCTGTGCGCGGTTAC	780	
Db	66447	CTGATGTGGGTGATCGATGAGATCAATCGTGGGAGCTGGAGATTCCTGTGCGCTCGCTAT	66506	
Qy	781	CAGTCAGATCCACCCCGGTTTATCGCTTATATACCGAAAAGATAGCTGCGCTCAAG	840	
Db	66507	CAGTCAGATCCACCCCGGTTTATCGCTTATATACCGAAAAGATAGCTGCGCTTAA	66566	
Qy	841	GTACAGGTCTGTATCAAACTCGCTCAGCGATTTATTTTGTGAGGTGCGTAAATTTTTCAG	900	
Db	66567	GTGAGGTGATGATTAACGGGTTAACTGACTATTTTGTGATGTCGACACCTGTTTTCAG	66626	
Qy	901	GAGATGACGGCGCGGGAAGAGAGTAA	930	
Db	66627	GAGATGACGGCGCGGGAAGAGAGTAA	66656	
RESULT 11				
AE008855				
LOCUS	AE008855	21252 bp	DNA	linear BCT 23-APR-2003

DEFINITION Salmonella typhimurium LT2, section 159 of 220 of the complete genome.

ACCESSION AE008855 AE006468

VERSION AE008855.1 GI:16421920

KEYWORDS

SOURCE

ORGANISM

Salmonella typhimurium LT2

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE

AUTHORS

McClalland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, B., Sun, H., Flores, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., and Wilson, R.K.

TITLE

Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2

JOURNAL

MEDLINE

21534948

PUBMED

11677609

REFERENCE

2 (bases 1 to 21252)

AUTHORS

The Salmonella typhimurium Genome Sequencing Project

CONSTRM

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA

COMMENT

Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs: GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and GeneMark; <http://opal.biology.gatech.edu/GeneMark/>

EC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and RegionDB; http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one mi3 subclone.

FEATURES

source

1..21252

/organism="Salmonella typhimurium LT2"

/mol_type="genomic DNA"

/strain="LT2; SGSC 1412; ATCC 700720"

/db_xref="ATCC:700720"

/db_xref="taxon:99287"

/note="LT2"

207..482

/gene="yhcn"

/note="synonym: STM3361"

207..212

/gene="yhcn"

/note="putative RBS for yhcN; RegulonDB: STMSIH003374"

219..482

/gene="yhcn"

/note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 75% identity in aa 18 - 104"

/codon_start=1

/transl_table=11

/product="putative outer membrane protein"

/protein_id="AAL22230.1"

/db_xref="GI:16421921"

/translation="MKIKTTVATLSILSVLFGAFAPFAEPIFAEQNQREAIQSVSVSA IGSSPMDMNAUUSKKADEQAQATYHTEARSQSNWHTAELYK"

579..852

/gene="STM3362"

579..584

/gene="STM3362"

/note="putative RBS for STM3362; RegulonDB: STMSIH003375"

586..852

/gene="STM3362"

/note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 53% identity in aa 18 - 104"

/codon_start=1

/transl_table=11

/product="putative periplasmic protein"

/protein_id="AAL22231.1"

/db_xref="GI:16421922"

/translation="MKTIVIIASLGLATLISFGANAHVQNAEQANLQPMGTISVS QIGSTPMDMRQRIKAKAGANSYRIIEIKEDNWHATAELYK"

complement(912..1197)

/gene="yhco"

/note="synonym: STM3363"

complement(912..1184)

/gene="yhco"

/note="similar to E. coli orf, hypothetical protein (AAC76271.1); Blastp hit to AAC76271.1 (90 aa), 76% identity in aa 1 - 90"

/codon_start=1

/transl_table=11

/product="putative cytoplasmic protein"

/protein_id="AAL22232.1"

/db_xref="GI:16421923"

/translation="MNVVTDFNDIKQSDFYREPTOTFGLASEKVDLTLDAVMS DILPLLEIEFVHLPDKLRRYGALILLFDEAEELGRLENVRH"

complement(1192..1197)

/gene="yhco"

/note="putative RBS for yhcO; RegulonDB: STMSIH003376"

complement(1356..3323)

/gene="yhcp"

/note="synonym: STM3364"

complement(1356..3323)

/gene="yhcp"

/note="similar to E. coli orf, hypothetical protein (AAC76272.1); Blastp hit to AAC76272.1 (655 aa), 92% identity in aa 1 - 655"

/codon_start=1

/transl_table=11

/product="putative inner membrane protein"

/protein_id="AAL22233.1"

/db_xref="GI:16421924"

/translation="MGIFSIANQHIRPAVKACAIVALFICGFQLETPRWAVLTAA IVAAGPAPAGGEPYSGAIRYRGMLRIIGTFCICAAALIIISIRAPMLILVCCVW AGCTWISLVRIENSYAGLSYTAIIIVITOTEPILTPQALRGSEIVIGICGA ILADLLFSRSIKQEVDELSDLLVAQOLMOLCIKHGDSREVDNAGDWLVRTALE GHRNLSNMRSSWRFRANRLKALNTLSLTIQSCETYLIIQNTRELITDTRELPEF PVEVDVHRQKRRRVIVTGERETPVLTLSVWGAATRYLLKRGVLSNTRISATE EILQGPVVKVIESAERHAMVNFRTLLSCILGTLFWLTGWTSGNAGWVIAVVT LAMRLNPRMVIDFTYGTAAALPLGLLYFLVIIPNTQSMLLCLSLAVLGFFIGIE VOKRLLSGMALATINIIIVLDNPMTFHSQFLDSALGQIVGCMIAFVILLVRDQSK DRTGRVLLNQFVSAASVAMTNVVRKENRLPALYQQLFLLNKPGLDKPKRLATM IIAHOSLRDAPIVNEDLSVFHQRERADHVISAGSDDKRRRYFQLLDELDTIYQEK LRIWEAPQVTFVKRLTGMHKYQVALTDS"

complement(329..4270)

/gene="yhcn"

/note="synonym: STM3365"

complement(3329..4261)

/gene="yhcn"

/note="similar to E. coli putative membrane protein (AAC76273.1); Blastp hit to AAC76273.1 (310 aa), 92% identity in aa 1 - 310"

/codon_start=1

/transl_table=11

```

/product="putative membrane located multidrug resistance protein"
/protein_id="AAL22234.1"
/db_xref="GI:16421925"
/translaton="MKTLTRKLSRTAITLVILAFIAFRAWVYTESPWTDRARFS
EKROAGRNRLGVQARSREILDOANNVLOTVLHQLAKAQATFDLAKDLERTVAP
ADGWYNLNVAGFEITRGSTALVYKNSFYQAYMEETKLEGVFGFRAEITPLGGS
NRLVKGTVSAAQGTWNASSTSDAKMGATVDSLEWVLAQRVFPVIRILDEQQGNLW
AGTTAVITTCQKRDASQDSFFPKLAHLREFG"
RBS
/gene="yhqQ"
/complement(4265..4270)
/gene="yhqQ"
/notes="putative RBS for yhcQ; RegulonDB:STWMS1H003377"
complement(4269..4481)
/gene="yhcR"
/gene="yhcR"
/notes="synonym: STM3366"
complement(4269..4472)
/gene="yhcR"
/notes="similar to E. coli orf, hypothetical protein (AAC76274.1); Blastp hit to AAC76274.1 (90 aa), 98% identity in aa 24 - 90"
/codon_start=1
/transl_table=11
/product="putative inner membrane protein"
/protein_id="AAL22235.1"
/db_xref="GI:16421926"
/translaton="WSLPFPVIVRGLSPFPIFFELLLSLAI FWLVRMLVPTGIYDFV
WHPALENTALYCCFLYILSRFLV"
RBS
/complement(4476..4481)
/gene="yhcR"
/notes="putative RBS for yhcR; RegulonDB:STWMS1H003378"
4654..5583
/gene="yhcS"
/notes="synonym: STM3367"
4654..5583
/gene="yhcS"
/notes="similar to E. coli putative transcriptional regulator LYSR-type (AAC76275.1); Blastp hit to AAC76275.1 (309 aa), 95% identity in aa 1 - 309"
/codon_start=1
/transl_table=11
/product="putative lysR family transcriptional regulator"
/protein_id="AAL22236.1"
/db_xref="GI:16421927"
/translaton="MESLRKMSVFAKVFSGSTAAARQLQMSVSSISQTVAKLEDEL
QVLLNSTRSISGITEAGKIYYOCRRMLHEVDQVHEQYVAFNNTIGTLRIGCSST
ACNVLAGTLAKLKYVGLAVNLTGYPADPLADGLVAVRVGALQDSSLRSRLGA
PMPVCAKPKVLAQGVYKPEADLSHSLWLEYSVRPNDFELTAPGEGITRLLPGKRF
VTNDDQVNLWALTDFGVDVAHLFGVMHGRGKEK"
RBS
/complement(5706..7151)
/gene="tldB"
/notes="synonym: STM3368"
71..64; Score 666; DB 1; Length 21252;
St. Local Similarity 82.3%; Pred. No. 8.3e-162; Indels 0; Gaps 0;
Matches 765; Conservative 0; Mismatches 165;
1 ATGGAACGACTAAAACGCATGTCGGTGTGTTGCCAAAGTAGTTCGAATTGGCTCTTTTACC 60
4654 ATGGAACGATTAAAACGAAATGTCGGTATTCGCCAAAGTTGTCGAGTTTGCTCTTTACC 4713
61 GCCGCCGCCACACAGCTACAGATCAGCGTTTCGTCCTCATCATCAGTCAGCGGTATCAAACTG 120
4714 GCCCAGCCGACAGCTGCAATGAGCGGTATCGTCCATCATGCCAGAGCTGTGGCGCAACTG 4773
121 GAAGATGAGTTGCAGGTAAAGCTGTWATAACCGTAGCACACGCAGCATTTGCCCTGCACGAA 180
4774 GAAGATGAGCTTCAGGTTAAATTCCTCAAACGCAGCACACGCAGCATTTGGGCTCACGAA 4833
181 GCCGCTAGAAATTTACTACCGAGGCTGCCGTCGTATGCTTATCATGAAGTCGAGGATGTCAT 240
4834 GCTGGAAAAATCTATTATCAGGGCTGTCGTCGTATGCTGATGAAGTCGAGGATGTTTCA 4893

```

Matches	733;	Conservative	0;	Mismatches	186;	Indels	0;	Gaps	0
Qy	1	ATGCAACGACTAAACAGCATGTCGGTGTGTGTCGCAAAAGTAGTTGAAATTTGGCTCTTTTACC	60						
Db	4	ATGCAACGATTAAACAGCATGTCGGTGTGTGTCGCAAAAGTAGTTGAAATTTAGTTCATTCACC	63						
Qy	61	GCGCGCCGACAGCAGCTACAGATGAGCGTTTCGTTCATCAGTCAGTACAGAGGGTATCAAAACTG	120						
Db	64	GCAGCGCCCGCCAGCTTCAGATGAGCGTTTCCTCGATCAGTCAGTCAGACCCGTCGCCAACTG	123						
Qy	121	GAAGATGAGTTCAGAGTAAAGCTGTTAAACCGTAGCACAGCAGCATTTGGCTGTACCGAA	180						
Db	124	GAAGATGAGCTCAGGTTAAAGCTGCTCAACCGCAGCACCCGACAGCCTTGGCTTAAGTAA	183						
Qy	181	GCGGTAGAAATTTACTACAGGGCTGCGGTGCTATGCTTCATGAAGTCAGGATGTTCAAT	240						
Db	184	GCAGGTAAAAATCTATTACAGGGCTGCGTAGGATGCTGTTTGAAGCGCAGGATGTCAT	243						
Qy	241	GAGCAACTGTATGCCTTCAATAACACCCCATCGGAGCTACGCAATTTGGCTGTGTTCTTCA	300						
Db	244	GAGCAATTTACGCAATTTAAACAATACCCCCATCGGAGCTGCGCATCGCTGCTCTTCA	303						
Qy	301	ACTATGGCAAAATGTTCTCGCGGGCTGACGCCAAAATGCTGAAGAATACCCAGCT	360						
Db	304	ACTATGGCACAGAAATGTTCTCGCGCGCATACGCGGAAATGCTCAAGATACCCCGG	363						
Qy	361	TTGAGCGTCAATTCGGTTACCGGAATTCAGCCCCGACCTGATTCGCCAGCGTCTGGAT	420						
Db	364	CTCTCTGTTAATCTGGTCACCGCATACCTCGCGCAGACCTGATGCCAGCGCTTGAC	423						
Qy	421	GTGGTATCCGGCTCGGCGGCTTGAGAGATTCCAGCCTGTTTCCCGCGCTCTGGCGCG	480						
Db	424	GTGGTGATCCGGTGGGCGCCCTCGAGGATTCAGACCTTTTTTCCCGCGCTCTGGGCTCA	483						
Qy	481	ATGCCAAATGGTGGTGGCGCGCAAAAGCTATCTCACAAATACGGCATACCGGAAAAA	540						
Db	484	ATGCCGATGGTATTATGCGCGCTATAAGCTATCTGGCGAGGCTGGCAATCCAGAGAA	543						
Qy	541	CCGCGGATTTGAGTAGTCATTATGCTTGAAATACAGCGTGGCGGCCGACAAATGAATTT	600						
Db	544	CCGCGCGACTTTGCCGACATGCTGGCTGGAAATATAGCTGGCGCGCGACAAATGAGTTT	603						
Qy	601	GACTGATCGACCGGAGGAGATCTGACTCGCCTGATCCCAAGAGGAGATTTGTGACT	660						
Db	604	GTCATTATCGCGCGGAAGGAGATCTGACCGCGCTGACCCCGCAGGGCGCTTCGTGACC	663						
Qy	661	AATGATCCGATGACCTGGTGGCTGGCTGACGCGGGTCCCGGGATCGCTACGTGGCG	720						
Db	664	AACGACCTATAGCGCTGGTGGCTGGCTGACCGCGGGGTGGGTATCGCTATGTACCC	723						
Qy	721	CTGATGTGGGTGATCAACAGATCAATGTGGGAGCTGGAGATCTCTGTCTGGCGGTTAC	780						
Db	724	TTAATGTGGCGATCGAAGAGATCAACCGCGGTGAACTTGAGATCTCTGTGCGCCAGCTAT	783						
Qy	781	CAGTCAGATCCACGCCCGGTTTATCGCTTATATACCGAAAAAGATTAAGCTCGCGCTGAG	840						
Db	784	CAGTCGATTCGCGGACCGGCTACGCCCTGTATACGGAAAAGGATAAATCTGGCGCTGAA	843						
Qy	841	GTACAGGTCGTGATCAACTCGCTGACGGAATTTTGTGAGTCCGGTAAATTTGTTTTCAG	900						
Db	844	GTCCAGGTCGTATCAACTATCTGACGGAGTATTTTGTGACGTGGCGAGATCTATCAG	903						
Qy	901	GAGATGCACGGCGCGGA	919						
Db	904	GGCATGCACGGCGCGGA	922						

DEC 17 1963

RESULT 13
BX950851 03/0

WPCOMMENT

Sequence split into 51 fragments LOCUS BX950851 Accession BX950851

Fragment Name
Fragment 1
Fragment 2
Fragment 3
Fragment 4
Fragment 5
Fragment 6
Fragment 7
Fragment 8
Fragment 9
Fragment 10
Fragment 11
Fragment 12
Fragment 13
Fragment 14
Fragment 15
Fragment 16
Fragment 17
Fragment 18
Fragment 19
Fragment 20
Fragment 21
Fragment 22
Fragment 23
Fragment 24
Fragment 25
Fragment 26
Fragment 27
Fragment 28
Fragment 29
Fragment 30
Fragment 31
Fragment 32
Fragment 33
Fragment 34
Fragment 35
Fragment 36
Fragment 37
Fragment 38
Fragment 39
Fragment 40
Fragment 41
Fragment 42
Fragment 43
Fragment 44
Fragment 45
Fragment 46
Fragment 47
Fragment 48
Fragment 49
Fragment 50
Fragment 51
Fragment 52
Fragment 53
Fragment 54
Fragment 55
Fragment 56
Fragment 57
Fragment 58
Fragment 59
Fragment 60
Fragment 61
Fragment 62
Fragment 63
Fragment 64
Fragment 65
Fragment 66
Fragment 67
Fragment 68
Fragment 69
Fragment 70
Fragment 71
Fragment 72
Fragment 73
Fragment 74
Fragment 75
Fragment 76
Fragment 77
Fragment 78
Fragment 79
Fragment 80
Fragment 81
Fragment 82
Fragment 83
Fragment 84
Fragment 85
Fragment 86
Fragment 87
Fragment 88
Fragment 89
Fragment 90
Fragment 91
Fragment 92
Fragment 93
Fragment 94
Fragment 95
Fragment 96
Fragment 97
Fragment 98
Fragment 99
Fragment 100

BX950851 00

BR950851_01	100001	210000
BR950851_02	200001	310000
BR950851_03	300001	410000
BR950851_04	400001	510000
BR950851_05	500001	610000
BR950851_06	600001	710000
BR950851_07	700001	810000
BR950851_08	800001	910000
BR950851_09	900001	1010000
BR950851_10	1000001	1110000
BR950851_11	1100001	1210000
BR950851_12	1200001	1310000
BR950851_13	1300001	1410000
BR950851_14	1400001	1510000
BR950851_15	1500001	1610000
BR950851_16	1600001	1710000
BR950851_17	1700001	1810000
BR950851_18	1800001	1910000
BR950851_19	1900001	2010000
BR950851_20	2000001	2110000
BR950851_21	2100001	2210000
BR950851_22	2200001	2310000
BR950851_23	2300001	2410000
BR950851_24	2400001	2510000
BR950851_25	2500001	2610000
BR950851_26	2600001	2710000
BR950851_27	2700001	2810000
BR950851_28	2800001	2910000
BR950851_29	2900001	3010000
BR950851_30	3000001	3110000
BR950851_31	3100001	3210000
BR950851_32	3200001	3310000
BR950851_33	3300001	3410000
BR950851_34	3400001	3510000
BR950851_35	3500001	3610000
BR950851_36	3600001	3710000
BR950851_37	3700001	3810000
BR950851_38	3800001	3910000
BR950851_39	3900001	4010000
BR950851_40	4000001	4110000
BR950851_41	4100001	4210000
BR950851_42	4200001	4310000
BR950851_43	4300001	4410000
BR950851_44	4400001	4510000
BR950851_45	4500001	4610000
BR950851_46	4600001	4710000
BR950851_47	4700001	4810000
BR950851_48	4800001	4910000
BR950851_49	4900001	5010000
BR950851_50	5000001	5064019

Overall Match	50 18:	Score	550:	DB 1:	Length
					110000:

Query Match	
Best [local] Similarity	76.3%
Pred. No.	1.3e-131:
Score	350, 25 1,
33.1%,	

Best Local Similarity 70.3%, Fred: NO, 1.36 137,
Matches 676: Conservative 0: Mismatches 210: Indels 0: Gaps 0:

1 ATCGAACGACCTAAACCCCATCTCCGCTGTTGCCAAAGTAGTTGAATTGGCTCTTTTACC

25429 ATGGGAAAGGACCTAAGACCTATGTCGGTGTTTGGCCCGCGGTGGTGGAAATGTGTGGTCTTTTATCC 25420

61 GCGGCGCGCGACAGCTACAGATCAGTCCGTTCCATCAGTCAGACGGTATC

[illegible]

180

25300

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
[
\
]
^
_
`
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x
y
z
{
|
}
~
_


```
Db 25239 GAACAGCTTTATGCTTCAACACACGCGGATCGGCACGCTACGATTGGTAGTTCGTCT 25180
QY 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGT 360
Db 25179 ACCATGGCCAGATGTTATTCACCATGATGCGCGATGCTGAAGATATCTCTGGT 25120
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGCGACCTGATTCGCGACGCTTGAT 420
Db 25119 TTGTCGCTCAATCTGTCACCGGCATTCGCGGTCCCGACTGATCGCGACGACTGGAT 25060
QY 421 GTGGTGATCCGGTGGCGCGCTGAGGATTCAGCGCTGTTTCCGCGCTGCGCGCG 480
Db 25059 ATCGTTATCCGGTGGTGACATCAGGATTCACCGCTGTTTCCGAAAGTTTAGGCTCG 25000
QY 481 ATGCCAATCGTGTGCGCCCGCAAGATCTCTCACACAATACGGCATACCGGAAAAA 540
Db 24999 ATGCCGATGGTGTGTCGCGGAAAGCTATCTGGCGCACACGGCACGCGGATTAAG 24940
QY 541 CCGCGGATTTGAGTAGTCAATTCATGCTGTTGATACAGCGGTGCGCGCGACATTAAT 600
Db 24939 CAGCGAGATATGTTGAAATTTTCCGCTGGGATACAGCGTGGCGCGCGACAGTGAAT 24880
QY 601 GAACGATCGCAGCGGAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
Db 24879 GAACGATTCACCGGAGGATTTGACCCCGCTCACACACACGCGGCTTTGTCAAC 24820
QY 661 ATGATCCGATGACGCTGTGCGTGTGCTGACGCGGGTGGCGGATCGCTACGTGCG 720
Db 24819 AACGATCCGCAACGCTGTGCTGGCTCAAGATCGGAGCGGTATCGCGTATATCCG 24760
QY 721 CTGATGCGGTGATCAACAGATCAATCTGGGGAGCTGGAGATCTGTCGCGGTTAC 780
Db 24759 CTCATGTGATTTGGACGAATCAATCTGGGAAATCGAGATCTATTCAACGCTAC 24700
QY 781 CAGTCAGATCCAGCGCGCTTTATCGTTPATACCGAAAAAGATAAGCTGCGCTGAAG 840
Db 24699 CACTCGATCCGCGCGCTACGCGCTTACCGCTTACCGCAAGACAACTGCGCGTAAA 24640
QY 841 GTACAGTCTGATCACTCGCTGCGGATTTTGTGAGTCTG 886
Db 24639 GTACAGTCTGCTAATTAATATGACGGAATCTTCAAGAAAGCTCG 24594

RESULT 14
ECOTLDD1/c 2788 bp DNA linear BCT 14-APR-2000
LOCUS Escherichia coli DNA for Tidd protein, complete cds.
DEFINITION D4451
VERSION D4451.1 Gi:1732436
KEYWORDS Tidd protein.
SOURCE Escherichia coli
ORGANISM Escherichia coli
REFERENCE 1 (sites)
AUTHORS Murayama,N., Shimizu,H., Takiguchi,S., Baba,Y., Amino,H.,
Horiuchi,T., Sekimizu,K. and Miki,T.
TITLE Evidence for involvement of Escherichia coli genes pmBA, csrA and a
previously unrecognized gene tidd, in the control of DNA gyrase by
letD (ccdB) of sex factor F
J. Mol. Biol. 256 (3), 483-502 (1996)
MEDLINE 96177756
PUBMED 8604133
REFERENCE 2 (bases 1 to 2788)
AUTHORS Miki,T.
TITLE Direct Submission
JOURNAL Submitted (24-DEC-1994) Takeyoshi Miki, Kyushu University, Faculty
of Pharmaceutical Sciences; 3-1-1, Maidashi, Higashi-ku, Fukuoka,
Fukuoka 812, Japan (E-mail:miki@bisei.phar.kyushu-u.ac.jp,
Tel:092-641-1151(ex.6187), Fax:092-632-6648)
LOCATION/Qualifiers
FEATURES
source 1..2788
```

```
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12"
/isolate="KP4714"
/db_xref="taxon:562"
/clone="pKP1846"
275..280
/evidence=not_experimental
297..618
/notes="91 bp direct repeat; each repeat contains two REP
(repetitive extragenic palindromic) sequences"
/citations=[1]
/rpt_type=tandem
/rpt_unit="345..435"
298..303
/evidence=not_experimental
306..344
/citations=[1]
/rpt_family="REP"
345..435
/citations=[1]
/rpt_type=inverted
436..526
/citations=[1]
/rpt_family="REP"
527..618
/citations=[1]
/rpt_type=inverted
618..650
/evidence=not_experimental
662..2107
/genes="tidd"
/function="Modulator of interaction between LetD protein
and DNA gyrase"
/codon_start=1
/transl_table=11
/product="Tidd protein"
/protein_id="BAA07913.1"
/db_xref="GI:1732437"
/translations="MSLNLVSEOLLAANGKHQDLFAILQLAERLDYGLYFOSSY
HESWLEDRIDKGSYNIDQGVRAISGEKTFAYADOISLALPQSAQAARTIVRD
SGDKVQTIGAVEHSPLYTSVDPLOSMREKDLILRDKVAREADKEVQVETASLS
GEYRDLAATDGLAADVRPLVRSVLVEDKRRGASGGGRFGFEFLADLD
GEYRADAAXEAVRMALVNLASVAPAGTPVVLGAGWPVLLHEAVHGLEDFNRR
GTSVFGQVGLVASLCTVDDGTVDGRSVAIDDEGTGGYNYLIENGILKGYMQ
DKLNARLMGTPGTNGRRSYAHLPMRMTNTYMLPKGTPOEIIESVEYGIYAPNFG
GGOVDITSGKFVFTSEAYLIENGKVTVPKGTALIGSGIETMQQISMVGNLKLNG
VGVCGEQSLPVGQPIKVDNLTVGTA"
2036
/genes="tidd50"
2036
/genes="tidd50"
/citation=[1]
/phenotype="Tolerant to LetD product growth inhibition"
/replace="t"

ORIGIN
Query Match 58.9%; Score 548; DB 1; Length 2788;
Best Local Similarity 100.0%; Pred. No. 3.2e-131;
Matches 548; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 383 GAATTCAGCCCCCGACCTGATTGCCGAGCTGTGGATGTGGATCCGCGCGCGCT 442
Db 2788 GAATTCAGCCCCCGACCTGATTGCCGAGCTGTGGATGTGGATCCGCGCGCGCT 2729
QY 443 TGCAGGATTCAGCCCTGTTTCCCGCGCTGTGGCGCGCATGCCAATGGTGGTGTGGCGCG 502
Db 2728 TGCAGGATTCAGCCCTGTTTCCCGCGCTGTGGCGCGCATGCCAATGGTGGTGTGGCGCG 2669
```

QY 503 CGAAAGCTATCTCACAAATACGGCATACCGGAAAAACCCGCCGATTTGAGTAGTCATT 562
 Db 2668 CGAAAGCTATCTCACAAATACGGCATACCGGAAAAACCCGCCGATTTGAGTAGTCATT 2609
 QY 563 CATGCTTGAATACAGCGTGGCGCCGACATGAATTTGAATGATCGACCGAAGGGA 622
 Db 2608 CATGCTTGAATACAGCGTGGCGCCGACATGAATTTGAATGATCGACCGAAGGGA 2549
 QY 623 TCTCGACTCGCTGATCCACAAAGAAATTTGTGACTAAATGATCCGATGACGCTGTGTC 682
 Db 2548 TCTCGACTCGCTGATCCACAAAGAAATTTGTGACTAAATGATCGACCGTGTGTC 2489
 QY 683 GTGCGTGAACGCGGCTGCGGGATCGCTAGTCCCGCTGATGTGGTGATCAACGAGA 742
 Db 2488 GTGCGTGAACGCGGCTGCGGGATCGCTAGTCCCGCTGATGTGGTGATCAACGAGA 2429
 QY 743 TCAATCGTGGGAGCTGGAGATCCTGCTGCGCGTTACAGTCAGATCCAGCCCGGTTT 802
 Db 2428 TCAATCGTGGGAGCTGGAGATCCTGCTGCGCGTTACAGTCAGATCCAGCCCGGTTT 2369
 QY 803 ATGCGTTATATACGAAAGAAATAGTCCCGCTGAAGGTACAGTCTGTATCAACTGCG 862
 Db 2368 ATGCGTTATATACGAAAGAAATAGTCCCGCTGAAGGTACAGTCTGTATCAACTGCG 2309
 QY 863 TCACGAGTATTTGTTGAGGTCGGTAAATTTGTTTACGAGATGACGCGGCGCGGAAAG 922
 Db 2308 TCACGAGTATTTGTTGAGGTCGGTAAATTTGTTTACGAGATGACGCGGCGCGGAAAG 2249
 QY 923 AGAAGTAA 930
 Db 2248 AGAAGTAA 2241

RESULT 15
 AB077386 1512 bp DNA linear BCT 02-FEB-2002
 LOCUS Serratia marcescens tdrA gene for putative temperature-dependent
 DEFINITION regulator A, complete cds.
 ACCESSION AB077386
 VERSION 1
 KEYWORDS GI:18478992
 SOURCE Serratia marcescens
 ORGANISM Serratia marcescens
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Serratia.

REFERENCE 1
 AUTHORS Baba, M., Midorikawa, Y., Nakagawa, Y., Fujita, M. and Matsuyama, T.
 TITLE Serratia marcescens and Escherichia coli genes controlling
 temperature-dependent production of structurally unrelated
 secondary metabolites such as prodigiosin and serrawettin
 UNpublished
 JOURNAL 2 (bases 1 to 1512)
 NAKAGAWA, Y., BABA, M. and MATSUYAMA, T.
 DIRECT SUBMISSION
 SUBMITTED (09-JAN-2002) Yoji Nakagawa, Niigata University, Faculty
 of Agriculture, 8050, Ikarashi 2-no-cho, Niigata City, Niigata
 950-2181, Japan (E-mail:nakagawa@agr.niigata-u.ac.jp,
 Tel:81-25-262-6360(ex.6360), Fax:81-25-262-6360)

FEATURES
 source 1..1512
 /organism="Serratia marcescens"
 /mol_type="genomic DNA"
 /db_xref="taxon:615"
 301..1212
 /gene="tdrA"
 301..1212
 /gene="tdrA"
 /codon_start=1
 /transl_table=1
 /product="putative temperature-dependent regulator A"
 /protein_id="BAB84544.1"
 /db_xref="GI:18478993"
 /translation="MERLRMSVFAKVVFEFGSFTAAARQLGMSVSSISQTVSKLENEL"

ORIGIN
 Query Match 56.8%; Score 528.2; DB 1; Length 1512;
 Best Local Similarity 74.1%; Pred. No. 4.3e-126;
 Matches 668; Conservative 0; Mismatches 233; Indels 0; Gaps 0;

QY 1 ATGCGACGACTAAACGCGATGTCGGTGTGTCGAAAGTAGTTGAATTTGGCTCTTTTACC 60
 Db 301 ATGCGAAAGATTAAACGGAATGTCGGTATTGCGCAAGTGGTTGAGTTCGGTCTGTTACC 360
 QY 61 GCGCGCCGACAGACTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 120
 Db 361 GCGCGCCGCGCAACTCGGCATGAGCGTTTCATCGATCAGCAAAACCGTCTCGAACTG 420
 QY 121 GAAGATGAGTTCAGGTAAGCTGTTAAACCGTAGCACACGAGGATTCGCCCTGACCGAA 180
 Db 421 GAAATGAGTTCAGGTAAGCTGTTAAACCGTAGCACACGAGGATTCGCCCTGACCGAG 480
 QY 181 GCGGTGAGATTTACTTACAGGGCTGCCGTGCTATGCTTTCATGAAAGTCAGGATGTTTCA 240
 Db 481 GCGGCAAAATCTACTTATCAGGGCTGCCGACGATGCTCAGGAAAGTCAGCGAAGTGCAT 540
 QY 241 GAGCACTGATGCTTCAATTAACACCCCATCGGACGCTACGATTCGGTCTGTTCTTCA 300
 Db 541 GAAACGCTGATGCTGTTTAAACAATACGCGCGCGACGCTCGCATCGGCACTCATCC 600
 QY 301 ACTATGCGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGT 360
 Db 601 ACCATGCGCAAAACGTCGTGCGGAACATGACCGCGGAGATGATGAAGAGTACCGGSC 660
 QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCGGACCTGATTCGCGACGGTCTGGAT 420
 Db 661 CTGACGCTCAACCTGTTAAACCGGCAATTCGCGCGCGGATCTGATCACCGACGGTGGAT 720
 QY 421 GTGGTATCGCGTGGCGGCTGAGGATTCAGGCTGTTTTCGCGCGCTGTCGCGCG 480
 Db 721 CTGGTATCGGACAGAGGCTGCGAGGACTCAGGCTGTTCTTCGCGCGCTGCGGCG 780
 QY 481 ATCCCAATGCTGTGTGCGCGCAAAAGCTATCTACACAATACGGCATACCGGAAAA 540
 Db 781 ATCCGATGCTGTGTGCGCGCAAAAGCTACCTGATCCAGCAGGCAACCGCGAGAAG 840
 QY 541 CCGCGCGATTTGAGTAGTCAATTCATGCTTGAATACAGCGTGGCGCGCCGACATGAAT 600
 Db 841 CCGAGCGATGCTGCTCAACTTCTCTGCTGAGTACAGCGTGGCGCGCCGACAGGATTC 900
 QY 601 GAACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGACT 660
 Db 901 GAACTGATCGCGCGGAAGGATCACACGCGATTTCCCGCAGGGGCGTTCGTACC 960
 QY 661 AATGATCGGATGAGCTGTGTGCTGCTGAGCGGGTGGCGGATTCGCTACGTCGCG 720
 Db 961 AACGACTCTCCACCAATGATCCGTTGGCTGAAAAACGCGCGCGGATTCGCTACGCGCG 1020
 QY 721 CTGATGTTGGTATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTGTCGCGGCTTAC 780
 Db 1021 CTGATGTTGGTATCAACGAGATCAACCGCGGAGATCGAGATCTCTGTTCAAAGCTAC 1080
 QY 781 CACTCAGATCCAGCCCGGTTATGCGTTATATACCGAAAAAGATGAGTTCGCGCTGAAG 840
 Db 1081 CATTGATCCGCGGCGGATCTACGCGCTCTATACCGAGAAGCAAACTGCGCTGAAG 1140
 QY 841 GTACAGTCTGATCAACTCGCTGAGGATTTATTTGTTGAGTCGGTAAATTTGTTTAC 900
 Db 1141 GTACAGTCTGATCAATTTATTTACGACTATTTTCGACGCTAGCGCGGTCTATCAG 1200
 QY 901 G 901

QVLLNRSTRSLTEAGKIYQGRMLQEVSEVHEQLYAFNNTPAGTLRIGSSSTM
 AQNVLANMTAEMMKYPLGLTVNLVTGIPAPDLITDGLDVLRTALQDSSLFRRLLGQ
 MPMVCAAKSYLIQHTPKPDMVNFVSPFSEFELMSPEGTTTISPGFR
 VTNDSTWIRLWLNKGAGIAYAPLMWVIBEIKRKEIEILFKSVHSDPRPIYALYTEKDK
 LPLKQVCVINYLTDFERVAAYQGYR"

Db 1201 G 1201

Search completed: December 19, 2004, 01:09:49
Job time : 4562.99 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:46:30 ; Search time 610.159 Seconds
(without alignments)
8001.134 Million cell updates/sec

Title: US-10-759-889-1

Perfect score: 930

Sequence: 1 atgggaacgactaaacgcgat.....ggcgggaaagagaagtaa 930

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	853.6	91.8	1742	5	AAS89852
2	621.4	66.8	933	11	ABD00728
3	407.6	43.8	1120	5	AAS93346
4	407.6	43.8	1575	5	AAS94105
5	369	39.7	750	5	AAS89851
6	122	13.1	948	11	ABD02039
7	121.8	13.1	1026	11	ABD06582
8	112.6	12.1	927	11	ACH96283
9	112.4	12.1	939	8	ACA24126
10	110.4	11.9	957	11	ABD03348
11	110.4	11.9	1047	11	ABD05418
12	110.4	11.9	1089	11	ABD05289
13	109.6	11.8	945	8	ACA44271
14	108.6	11.7	915	11	ACH96991
15	108	11.6	912	11	ABD02170
16	107.6	11.6	1017	11	ABD13371
17	107.6	11.6	1104	11	ABD13302
18	106.4	11.4	1029	11	ABD15082
19	106.4	11.4	1071	11	ABD14855
20	106.4	11.4	1809	11	ABD15283
21	105.6	11.4	936	8	ACA25126

22	105	11.3	900	3	AAS23355
23	103.6	11.1	909	10	ABZ41110
24	103.6	11.1	912	3	AAS23353
25	103.4	11.1	888	10	ADC19769
26	103.4	11.1	900	3	AAS23354
27	103.4	11.1	1132	10	ADC19786
28	103.4	11.1	47475	3	AAA81465
29	103.4	11.1	110000	3	AAA81489
30	103.4	11.1	349980	3	AAF21612
31	103.2	11.1	933	11	ACH94659
32	102.8	11.1	777	11	ABD06646
33	102.2	11.0	591	10	ADH82126
34	101.4	10.9	915	11	ABD16364
35	101.4	10.9	1011	11	ABD16178
36	100.4	10.8	1062	11	ABD10482
37	100.2	10.8	1008	11	ACH97551
38	100	10.8	100	8	ACD69195
39	100	10.8	100	8	ACD69193
40	100	10.8	100	8	ACD69194
41	99.8	10.7	918	8	ABT14599
42	99.2	10.7	72750	3	AAA81468
43	99.2	10.7	110000	3	AAA81490
44	99.2	10.7	349980	3	AAF21544
45	98.6	10.6	792	11	ABD13408

ALIGNMENTS

RESULT 1

AAS89852

ID AAS89852 standard; cDNA; 1742 BP.

AC AAS89852;

XX

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #25656.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; as.

XX Homo sapiens.

XX WO300175067-A2.

XX PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG325665.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensic, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 25656; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal

activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: the sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 91.8%; Score 853.6; DB 5; Length 1742;
Best Local Similarity 99.3%; Pred. NO. 2.3e-253;
Matches 878; Conservative 0; Mismatches 4; Indels 2; Gaps 2

Qy	1	ATGGAACGACTAAACCGCATGTCGGTGTTCCTCCAAAGTAGTTCGAAATTTGGCTCTTTTACC	60
Db	28	ATGGAACGACTAAACCGCATGTCGGTGTTCCTCCAAAGTAGTTCGAAATTTGGCTCTTTTACC	87
Qy	61	GCGCGCGCAGACACCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGGTATCAAAACTG	120
Db	88	GCGCGCGCAGACACCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGGTATCAAAACTG	147
Qy	121	GAAGATGAGTTCCAGGTAAGCTGTTAAACCGTAGACAACGAGCAATTTGGCTGACCGAA	180
Db	148	GAAGATGAGTTCCAGGTAAGCTGTTAAACCGTAGACAACGAGCAATTTGGCTGACCGAA	207
Qy	181	GCGGGTAGAATTTACTACACGGCTGCCGTCGTATGCTTCATCAAGTGCAGATGTTTC-A	239
Db	208	GCGGGTAGAATTTACTACACGGCTGCCGTCGTATGCTTCATCAAGTGCAGATGTTTC-A	267
Qy	240	TGAGCAACTGTATGCTCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTTCTTC	299
Db	268	TGAGCACTGTATGCTCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTTCTTC	327
Qy	300	AACTATGGCACAAAATGTTCTCGCGGCTGCAGAGCCAAATGCTGAAAGATACCCAG	359
Db	328	AACTATGGCACAAAATGTTCTCGCGGCTGCAGAGCCAAATGCTGAAAGATACCCAG	387
Qy	360	TTTGAGCGTCAATCTGGTTACGGAAATTCAGCCCCCGACCTGATTTGCCGAGGTCGGA	419
Db	388	TTTGAGCGTCAATCTGGTTACGGAAATTCAGCCCCCGACCTGATTTGCCGAGGTCGGA	447
Qy	420	TGTGGTGATCCGGTCGGCGGTTGCAGGATTCAGAGCTGTTTTCGCCGCGCTCGGGCGC	479
Db	448	TGTGGTGATCCGGTCGGCGGTTGCAGGATTCAGAGCTGTTTTCGCCGCGCTCGGGCGC	507
Qy	480	GATGCCAATGGTGGTGTGCGCGGAGAAAGCTATCTCACACAATACGGCAATACCGGAAA	539
Db	508	GATGCCAATGGTGGTGTGCGCGGAGAAAGCTATCTCACACAATACGGCAATACCGGAAA	567
Qy	540	ACCGCGGATTTGAGTAGTCAATCATGCTTTGAATACAGCGTGGCGCGACACATCAAT	599
Db	568	ACCGCGGATTTGAGTAGTCAATCATGCTTTGAATACAGCGTGGCGCGACACATCAAT	627
Qy	600	TGAACCTG-ATCGCACCGGAAGGATCTCGACTCGCTTATCCCAAGAGGAAGTTGTGA	658
Db	628	TGAACCTGATCGCACCGGAAGGATCTCGACTCGCTTATCCCAAGAGGAAGTTGTGA	687
Qy	659	CTAATGATCCGATGACGTCGTGCGCTGGCTGACGGCGGGTGCGGGATCGCCTACGTCG	718
Db	688	CTAATGATCCGATGACGTCGTGCGCTGGCTGACGGCGGGTGCGGGATCGCCTACGTCG	747
Qy	719	CGCTGATGTGGGTGATCAACGAGATCAATCTGTGGGAGCTGGAGATCCTCTCCCGCGTT	778
Db	748	CGCTGATGTGGGTGATCAACGAGATCAATCTGTGGGAGCTGGAGATCCTCTCCCGCGTT	807

Qy	779	ACCAGTCAGATCCAGCCGCGTTTATCGGTATATACCGNAAAGATAGCTGCCGCTGA	838
Db	808	ACCAGTCAGATCCAGCCGCGTTTATCGGTATATACCGNAAAGATAGCTGCCGCTGA	867
Qy	839	AGGTACAGGTCGTGATCAACTCGCTGACGGATTATTTGTTGAG	882
Db	868	AGGTACAGGTCGTGATCAACTCGCTGACGGATTATTTGTTGAG	911

RESULT 2

RESOL 2
ABD00728
ID ABD00728 standard; DNA; 933 BP.

XX
AC ABD00728;

DT 29-JUL-2004 (first entry)

XX
DE Klebsiella pneumoniae polynucleotide seqid 6503.

XX transcription regulatory element;
XX Recombinant expression vector;
KW Klebsiella pneumoniae protein;
KW Klebsiella pneumoniae; antibiotic;
XX Klebsiella pneumoniae;
OS Klebsiella pneumoniae.

XX
PN US6610836-B1.

XX
PD 26-AUG-2003.

XX
PF 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.
PR

XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Breton GL, Osborne M;

XX
DR WPI; 2003-895346/82.

DR p-PSDB; AB067157.
XX
PT New nucleic acid encoding a *Klebsiella pneumoniae* polypeptide, useful for
PT preparing a vaccine composition against *Klebsiella pneumoniae*.

PS Disclosure: SEO ID NO 6503; 932pp; English.

xx The invention describes a new isolated nucleic acid encoding a Klebsiella
xx pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention

Sequence 933 BP: 208 A; 263 C; 263 G; 199 T; 0 U; 0 Other;

Query Match	66.8%	Score 621.4;	DB 11;	Length 933;
Best Local Similarity	79.8%;	Pred. No. 1.4e-181;		
Matches 733;	Conservative	0;	Mismatches 186;	Indels 0;
Gaps	0			

Qy	1	ATGGAACGACTAAACCGCATGTCGGTGTGGCCAAAGTAGTTGAAATTTGGCTCTTTTACC	60
Db	4	ATGGAACGATTTAAACCGCATGTCGGTGTGGCCAAAGTGGTGAAATTAGTTCAATTCACC	63
Qy	61	GCGCCGCCAGACAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGGGTATCAAAACCTG	120
Db	64	GCAGCGGCCCGCAGCTTCAGATGAGCGTTTCCTCGATCAGTCAGACCGTGGCAAACTG	123
Qy	121	GAAGATGAGTGCAGGTAAACCTGTTAAACGCTAGCACGCGAGCATTTGGCTCTACCGAA	180
Db	124	GAAGATGAGCTCAGGTAAAGCTGCTCAACCGCAGCACCGCAGCCTTTGGCTTAACGTGA	183
Qy	181	GCGGGTAGAAATTTACTACACGGGCTGGCTGCTATGCTTCATGAAGTCGAGGATGTTCTAT	240

```
Db 184 GCAGGTAAATCTATTACACGGGCTGCGGTAGAGTCTGTTTGAAGCGCAGGATGTCAT 243
QY 241 GAGCAACTGTATGCTTCAATACACCCCATCGGACGCTACGCAATGGCTGTCTTCA 300
Db 244 GAGCAATTATACGATTTAAACATACCCCATCGGACGCTGGCATGCTGCTTCA 303
QY 301 ACTATGGCACAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAAGAAATACCGAGT 360
Db 304 ACTATGGCACAGAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAAGAAATACCGAGT 363
QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCACGCGCCCGACCTGANTGCCGAGGCTTGGAT 420
Db 364 CTCTCTGTTAACTGCTACCGGCATACCTGGCCAGACCTGATGCCGACGGCTTGAC 423
QY 421 GTGGTATCGGCTCGCGGCTTGCAGATTCAGGCTGTTTCCCGCGCTGCGCGCG 480
Db 424 GTGGTATCGGCTCGCGGCTTGCAGATTCAGGCTGTTTCCCGCGCTGCGCGCTCA 483
QY 481 ATGCCAATGTTGTTGCGCGGCTGACAGCAATCTCACAATACGGCATACCGGAAAA 540
Db 484 ATGCCAATGTTGTTGCGCGGCTGACAGCAATCTCACAATACGGCATACCGGAAAA 543
QY 541 CCGCGCGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Db 544 CCGCGCGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 603
QY 601 GAACTATCGCAGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 604 GTCATTATCGCGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 661 AATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 664 AAGACCCCTATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
QY 721 CTGATGTGGGTATCAACGAGATCAATCTGTTGGGAGCTGGAGATCTCTGCTGCTGCT 780
Db 724 TTAATGTGGGCGATCGAGAGATCAACCGCGGTGAATCTGATCTGCTGCTGCTGCTGCT 783
QY 781 CAGTCAGATCCACGCGGCTTTATGCTTATATACGAAAGATGAGCTGCGCTGAG 840
Db 784 CAGTCAGATCCACGCGGCTTTATGCTTATATACGAAAGATGAGCTGCGCTGAG 843
QY 841 GTACAGGTCGTATCAACTCGCTGACGATTTATTTGTTGAGGTCGTAATTTGTTTAC 900
Db 844 GTACAGGTCGTATCAACTCGCTGACGATTTATTTGTTGAGGTCGTAATTTGTTTAC 903
QY 901 GAGATGCAAGGCGCGGA 919
Db 904 GGCATGCAAGGCGCGGA 922

RESULT 3
ID AAS93346
XX AAS93346 standard; cDNA; 1120 BP.
AC AAS93346;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #29150.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
```

```
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PSDB; ABG29159.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 29150; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1120 BP; 271 A; 279 C; 323 G; 247 T; 0 U; 0 Other;
Query Match 43.8%; Score 407.6; DB 5; Length 1120;
Best Local Similarity 96.7%; Pred. No. 2.7e-115;
Matches 416; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 481 ATGCCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540
Db 1 ATGCCAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 60
QY 541 CCGCGCGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
Db 61 CCGCGCGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120
QY 601 GAACTATCGCAGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 121 GAACTATCGCAGCGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 661 AATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 181 AATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 721 CTGATGTGGGTATCAACGAGATCAATCTGTTGGGAGCTGGAGATCTCTGCTGCTGCTTAC 780
Db 241 CTGATGTGGGTATCAACGAGATCAATCTGTTGGGAGCTGGAGATCTCTGCTGCTGCTTAC 300
QY 781 CAGTCAGATCCACGCGGCTTTATGCTTATATACGAAAGATGAGCTGCGCTGAG 840
Db 301 CAGTCAGATCCACGCGGCTTTATGCTTATATACGAAAGATGAGCTGCGCTGAG 860
QY 841 GTACAGGTCGTATCAACTCGCTGACGATTTATTTGTTGAGGTCGTAATTTGTTTAC 900
Db 361 GTACAGGTCGTATCAACTCGCTGACGATTTATTTGTTGAGGTCGTAATTTGTTTAC 920
QY 901 GAGATGCAAG 910
```

```

Db      421 TTCGTGCACG 430
|||||
RESULT 4
AAS94105
ID AAS94105 standard; cDNA; 1575 BP.
XX
AC AAS94105;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #29909.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX DR P-PSDB; ABG29918.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 29909; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1575 BP; 402 A; 376 C; 459 G; 338 T; 0 U; 0 Other;

Query Match 43.8%; Score 407.6; DB 5; Length 1575;
Best Local Similarity 96.7%; Ident. No. 3.2e-115;
Matches 416; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

XX 481 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
DB 1 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 60

```


genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (III) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 39.7%; Score 369; DB 5; Length 750;
Best Local Similarity 82.7%; Pred. No. 2e-103;
Matches 472; Conservative 0; Mismatches 15; Indels 84; Gaps 1;

QY	340	ATGCTGAAGAATACCCAGGTTTGAGCGTCAATCTGGTTACCGAATTCAGACCCCGAC	399
Db	1	ATGCTGAAGAATACGACAGGTTTGAGCGTCAATCTGGTTACCGAATTCAGACCCCGAC	60
QY	400	CTGATTCGCACGCTCTGGATGTGGTGAATCGCGTTCGCGCGTTGCAGGATTCAGCGTG	459
Db	61	CTGATTCGCACGCTCTGGATGTGGTGAATCGCGTTCGCGCGTTGCAGGATTCAGCGTG	120
QY	460	TTTTCGCCGCGTCTGGCGCGCATGCCAATGTGTGTGCGCGCGGAAAAGCTATCTCACA	519
Db	121	TTTTCGCCGCGTCTGGCGCGCATGCCAATGTGTGT	155
QY	520	CAATACGGCATACCGGAAAAACCCGCCGATTTGATGTAGTCAATCATGGCTTGAATACAGC	579
Db	156	-----C	156
QY	580	GTGCGGCGCGACAAATGAATTTGAATCTGATCGCACCGGAAGGATCTCGACTCGCGTGAATC	639
Db	157	GTGCGGCGCGACAAATGAATTTGAATCTGATCGCACCGGAAGGATCTCGACTCGCGTGAATC	216
QY	640	CCACAGGAAGATTGTGACTTAATGATCCGATGACGCTGGTGGCGCTGGCTGACGCGCGGT	699
Db	217	CCACAGGAAGATTGTGACTTAATGATCCGATGACGCTGGTGGCGCTGGCTGACGCGCGGT	276
QY	700	GCGGGATCGCTACGTGCCGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTG	759
Db	277	GCGGGATCGCTACGTGCCGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTG	336
QY	760	GAGATCTGTGCTGCGCGTTACCACTCAGATCCACGCCCGGTTTATGGTTATATACCGAA	819
Db	337	GAGATCTGTGCTGCGCGTTACCACTCAGATCCACGCCCGGTTTATGGTTATATACCGAA	396
QY	820	AAAGATAAGCTCGCGCTGAAGGTACAGTCTGTGATCAACTCGCTGACGGATTATTTTGT	879
Db	397	AAAGATAAGCTCGCGCTGAAGGTACAGTCTGTGATCAACTCGCTGACGGATTATTTTGT	456
QY	880	GAGTCGGTAAATTTGTTTCAGGAGATCAACG	910
Db	457	GAGTCGGGCGAGGGACCGTTCTGTGCACG	487

RESULT 6
ABD02039
ID ABD02039 standard; DNA; 948 BP.
XX
XX
AC ABD02039;
XX
XX
29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #643.

XX KW KW XX OS XX PN PD PD XX PF XX PR PR PR PA XX PI XX DR DR XX PT PT PT XX PS XX CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC CC

0442

QY Db QY Db QY Db QY Db QY Db QY

Bacterial infection; gene; ds; *Pseudomonas aeruginosa* infection; antibacterial. .
Pseudomonas aeruginosa.
 US6551795-B1.
 22-APR-2003.
 18-FEB-1999; 98US-00252991.
 18-FEB-1998; 98US-0074788P.
 27-JUL-1998; 98US-0094190P.
 (GENO-) GENOME THERAPEUTICS CORP.
 Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 WPI; 2003-615309/58.
 P-PSDB; ABC69468.
 Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polyn-
 useful as molecular targets for diagnostics, prophylaxis and tre-
 pathological conditions resulting from bacterial infection.
 Disclosure: SEQ ID NO 643; 455pp; English.

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-*P. aeruginosa* drugs, as templates for recombinant production of *P. aeruginosa*-derived peptides or polypeptides, as target components for diagnosis and/or treatment of *P. aeruginosa*-caused infection, and in detection of *P. aeruginosa* sequences or other sequences of *Pseudomonas* species using biochip technology. Sequences ABD01397-ABD1967 represent *P. aeruginosa* polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

Sequence 948 BP; 156 A; 314 C; 330 G; 148 T; 0 U; 0 Other;

```

every Match      13.1%; Score 122; DB 11; Length 948;
at Local Similarity 46.9%; Pred. No. 7.4e-27;
Mismatches 417; Conservative 0; Mismatches 470; Indels 3; Gaps 1

```

1 ATGGAACGACTAAACCGCATCGGTGTTTGCCAAAGTAGTTGAAATTTGGCTCTTTTACC 60
40 ATGGATTGCGTGAAGGCATGCGNATCTTGCCACTGTGTTGACAAAGGCTCGATGGCA 99	
61 GCGCGGCCAGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAACTG 120	
100 GCGCGGCCAGAGGCTCGGCATGACCCCTTGGCAGTCAGCCAGCAGATCCGCAAGCTG 159	
121 GAAGATGAGTTCCAGGTAAAGCTGTTAAACCGTAGCACAGCGAGCATTGGCTGTACCCGAA 180	
160 GAGAGCCGCCCCAGTGTCATTTGCTGCATCGACACACCCGCCGCTGACCTGTCCGAG 219	
181 GCGGTAGAAATTTACTACAGGGTGCCTGTCGTATGCTTCATGAAGTCAGAGATGTTCA 240	
220 GCGCGAGGCGGCTTCTATCGCAGTTGCGCGCAGATGTGGCGATCGCCAGGAAGCCGAG 279	
241 GAGCACTGTATGCCCTTCAATAACACCCCATCGGACGCTACGCATTTGGCTGTCTTCA 300	
280 CGGCGGCTTGGCGAATGGCGGATTCGCGCGGTTGGTGAGTTGCGCTTCGCGGCGCGGTTG 339	
301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360

Db 340 GGGTTCTCGGCACGCTGATCACCCAGGCGCTGAAACCGTTGCTGTGAGAACACCCGCCAG 399
QY 361 TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCCCGACCTGATTTGCCGAGCGTCTGGAT 420
Db 400 TTGCGCTCGAGCTGTCTTCCAGGACGAGCGATCGATCTGGTGGCCGAGCGATCGAC 459
QY 421 GTGGTATCCGGTGGCGGCTGTCAGGATTCACCGCTTTTCCCGCGCTCGGGCGG 480
Db 460 CTGGCGATACGGTGGCGAATCTCCCGACTCCAGCTGTGGCGGCGACCTCGGGAT 519
QY 481 ATGCCAATGGTGTGGCCCGCGAAAGCTATCTCACAAATACGGCATACCGGAAAA 540
Db 520 TGGAGCAGCGTCTCTGCGCGCGCGGCGCTATTTCGCCCAACGGGCGCCCATCAACCGT 579
QY 541 CCGCGGATTTGAGTAGTCAATTCATGGCTT---GAATACGGTGGCGCGCGCAATGAA 597
Db 580 CCGGAGCATTTACCGAGTGCAGTGGATTTCCCTGGAACCTCGAACCATCTCAACAC 639
QY 598 TTGAACTGATCGCACCGGAGGATCTCGACTCGCTCTGATCCCAAGGAAGATTGTG 657
Db 640 CTGACGCTCAGCGGCGCGGCGGGAAGTCTGCAAACTGCGCTTGGAGCCACGGGTGCC 699
QY 658 ACTAATGATCCGATGACGCTGTGGCTGGCTGACGCGGGTGGCGGATCGCTTACGTG 717
Db 700 GCAACGGGATCTCGCGGTGCGTCACTTCACCTCGACGACTCGCGGTGTCTTACCA 759
QY 718 CGCTGATGTGGTGTATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTGCTCGCGGT 777
Db 760 CGCTGCGGAGGTGCGGACCGCTGACGACGCGGCGCTTGAACAGCTATTCGCGAG 819
QY 778 TACAGTACATCCACCGCGGTTTATGGTTATATACGAAAGATAGCTGCGCTG 837
Db 820 TGGAGGATTCGCGGCTCGGCTATCTATGCGTGTGACGCGCACGCGGCGGCGCGCC 879
QY 838 AAGGTACAGTCTGATCACTCACTCGTACGCGGATTTATTTGTTAGGTGCG 887
Db 880 AAGGTGAAGTCCCATGACGCGCTTGGCGAGCGTTTCGCCACGAGCGG 929

RESULT 7

ABD06582
ID ABD06582 standard; DNA; 1026 BP.
XX AC ABD06582;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polynucleotide #5186.
XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX KW antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX WPI; 2003-615309/58..
XX DR P-PSDB; AB073011.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 5186; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX

SQ Sequence 1026 BP; 152 A; 384 C; 339 G; 151 T; 0 U; 0 Other;

Query Match 13.1%; Score 121.8; DB 11; Length 1026;
Best Local Similarity 47.1%; Pred. No. 8.9e-27;
Matches 409; Conservative 0; Mismatches 457; Indels 3; Gaps 1;

QY 1 ATGGAACGACTAAACGCAATGTCGGTGTGTCGCAAGTAGTTGAAATTTGGCTCTTTTACC 60
Db 106 ATGAGCGGCTGACCCGCACTCGCGTCTTCGTGCAAGTCAATGACAGCGGACGACGACC 165
QY 61 GCCCGCGCAGACAGCTACAGATGAGCTTTCGTCCATCAGTCAGAGCTATCAAACTG 120
Db 166 GCCCGCGCAGGCTGTGGACATGTCGGGCGATGCTCGGCTACCTCCCGAGTTG 225
QY 121 GAAGATGAGTTGAGGTAAGCTTTAAACCGTAGCACACGAGCATTTGGCTTGACCGAA 180
Db 226 GAGGATGGGTGGCGCAGCGCTGTGTCACCGCAGCACCGCGCTGAGCTGACCGAC 285
QY 181 GCCGCTAGAAATTTACTACAGGCTCGCTGATGCTTCAATGAGTGAGGATGTTAT 240
Db 286 GCCCGCGCAGGCTGTGCGCAGTGTGCGCGAGATGCTCGCGTGGCGCAGCGCTGACG 345
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGATTTGGCTTCTTCA 300
Db 346 GCCACGGCGCAGACCGCGCGACAGCGCGCGGCACTCTGGCATCACGAGACGCTG 405
QY 301 ACTATGGCAAAATGTTTTCGCGGCTGACAGCCAAATGCTGAAAGATACCGAGT 360
Db 406 TCCTTGGCCAGGCTGGGTGACCCGTGCGCGCGCGCTTCTGAGCGCTACCGCGGC 465
QY 361 TTGAGCGTCAATCTGTTACCGGAATTTCCAGCCCCCGACCTGATTCCCGACGCTCTGGAT 420
Db 466 ACCGCCATCGACCTGAGGTCAACAGCAGGTGGTCAACTGGTGAGGAGCGATCGAC 525
QY 421 GTGGTATCGCGTGGCGGCTTTCAGGATTCAGGCTGTTTTCGCGCTGTGGCGCGG 480
Db 526 CTAGCCCTGGCATCGCCAAACAGCTCGACCCCAACCTGATCGCGCGCGCTTCGCGGAA 585
QY 481 ATGCCAATGGTGTGGCGCGAAAGCTATCTCACAAATACCGCATACCGGAAAA 540
Db 586 TGGCGCTCGGTGATCTGCGCGCGCGCGGACTACTCTGCGCGCGCACCGCGCGCGC 645
QY 541 CCGCGCGGATTTGAGTAGTCAATTCATGGCTTGAATACAGGCTGGCGCCCGACAAATGTT 600
Db 646 CCGGAGGACCTGGCGTGCACACACTGCTTACTTCTGCGCGCGAGT---CTC 702
QY 601 GAAGTATCGACCGGAGGATCTCGATCGCTGATCCCAAGAGAGATTGTGACT 660
Db 703 TGGCAGTTTCAGCGCGACGCGGACCGCGCTGCGGTGGGTGGGAGTCTCAGCGCC 762
QY 661 AATGATCCGATGACGCTGGTGGCTGGCTGACGCGCGGCTGCGGGATCGCTACGTCGCG 720
Db 763 AAGCATCCACCGTGTGCTGGAGCGCGCGCGCGCGGTGCGGGATCAGCCAGCAGCGG 822

PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU20256.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 11996; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 939 BP; 145 A; 319 C; 327 G; 148 T; 0 U; 0 Other;
 SQ

Query Match
 Best Local Similarity - 47.8%; Pred. No. 7e-24;
 Matches 394; Conservative 0; Mismatches 421; Indels 9; Gaps 2;

QY 1 ATGGAACGACTAAACGCGATGCGGTGTTGCAAGTAGTTGAAATTTGGCTCTTTTACC 60
 Db |||||
 QY 1 ATGGAATCGTTACAAACATGCGGTGTTGTCGCGGTGTCGACGCGGCGAGCTTTACC 60
 Db |||||
 QY 61 GCGCGCGCCAGACAGTACAGATGACGCTTTCGTCATCAGTCACAGCGTATCAAAATG 120
 Db |||||
 QY 61 GCGCGCGCCAGACAGTACAGATGACGCTTTCGTCATCAGTCACAGCGTATCAAAATG 120
 Db |||||
 QY 121 GAAGATGAGTTCAGTAAAGTGTAAACCTAGCACAGCAGATTGGCTGACCGAA 180
 Db |||||
 QY 121 GAGGCCACCTCGCGACCGGTCTTCCTGAACCGACGACGCGCGGATCGCTGACCGAG 180
 Db |||||
 QY 181 GCGGCTAGAAATTTACTACGAGGCTCGCGTCTGATGCTTTCATGAAGTCAGGATTTTCA 240
 Db |||||
 QY 181 GCGGCTAGAAATTTACTACGAGGCTCGCGTCTGATGCTTTCATGAAGTCAGGATTTTCA 240
 Db |||||

QY 241 GAGCAACTGTATGCTTCAATTAACACCCCATCGGAGCGCTACGATTGGCTGTCTTCA 300
 Db |||||
 QY 241 GCGGAGCGGGGACGCGACGCGCGCGCTCGCGCAAGCTGAAGTTCATTGCTTCAG 300
 Db |||||
 QY 301 ACTATGGCACAATAAGTCTCTCGCCGGGTGACAGCCAAATCTGAAAGAAATACCCAGT 360
 Db |||||
 QY 301 AGCTCGGCGACACATATCTGTGCGGCCATCGCGCTACGCGGCGCTATCGGAC 360
 Db |||||
 QY 361 TTGAGCGTCAATCTGTTTACCGAAATTCAGCCCCGACCTGATTCGCCAGCGTTCGAT 420
 Db |||||
 QY 361 GTGACGTCGAGCTGACGCTCGCGCAGCGGATGCGCGACCTGCTGACGAGGGTACGAC 420
 Db |||||
 QY 421 GTGCGTATCGGCTGCG---GCGGTTGAGGATTCAGCCTGTTTCCGCGCTGTGGGC 477
 Db |||||
 QY 421 GTGCGATCGTGTGCGCGCGACCTGCGCGATTCGGGCTCGTGTGCGAGCGGTTCGGC 480
 Db |||||
 QY 478 GCGATGCCAATGTGTGTCGCGCGGAAAGCTATCTCACAAATACGCGATACCGAA 537
 Db |||||
 QY 481 GAGAGCTACAGGCTGCTGCGCATCGCGGGCTACGTCGAGTCGCGACGCGCTGCGCAG 540
 Db |||||
 QY 538 AAACCCCGCATTTGAGTAGTCAATTCATGCTTGAATACAGCTGCGG-----CCGAC 591
 Db |||||
 QY 541 CGCGCGCGGATCTCGCAGCAGCAGCTGTCCTCGGGATGTCGCGCGGGCTTTCAC TTC 600
 Db |||||
 QY 592 AATGAATTTGAATGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGAAAGA 651
 Db |||||
 QY 601 GAGCAATGCGCTGCGCGGGCGGACGCGAGGTCGTGCCGATCACGCGCGCGCG 660
 Db |||||
 QY 652 TTTGTGACTAATGATCCGATGACGTGTCGCTGCGCTGACGCGCGGTGTCGCGGATCGCC 711
 Db |||||
 QY 661 TTTGCGGTGAACGTCGCGAGGCGCTCGCGGTGCGCGGTGCGGAGGGATGCGCGCTGCGC 720
 Db |||||
 QY 712 TACGTCGCGCTGATGCGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTG 771
 Db |||||
 QY 721 GGGCTCGCGCTTATTCGCGGATCGCTGCTGCTGCGTAGCGGACATCTGTCGCGGTG 780
 Db |||||
 QY 772 CCGCGTTTACGATCAGATCCAGCCCGCGTTTATGCGTTATATAC 815
 Db |||||
 QY 781 CCCGAGTACCGTTCGACGATGATGAACATCTACGCGCTGTATCC 824
 Db |||||

RESULT 10
 ABD05348
 ID ABD05348 standard; DNA; 957 BP.
 AC ABD05348;
 XX
 XX 29-JUL-2004 (first entry)
 DT
 XX Pseudomonas aeruginosa polynucleotide #3952.
 DE Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 XX antibacterial.
 KW Pseudomonas aeruginosa.
 XX US6551795-B1.
 XX
 XX 22-APR-2003.
 PD
 XX 18-FEB-1999; 99US-00252991.
 PF
 XX 18-FEB-1998; 98US-0074788P.
 PR
 XX 27-JUL-1998; 98US-0094190P.
 PR
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 PI WPI; 2003-615309/58.
 XX P-PSDB; ABO1777.
 XX
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT

PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 3952; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biotech technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

SQ Sequence 957 BP; 145 A; 341 C; 327 G; 144 T; 0 U; 0 Other;

Query Match 11.9%; Score 110.4; DB 11; Length 957;

Best Local Similarity 50.0%; Pred. No. 2.9e-23;

Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCTATGCGTGTGTCGCAAGTAGTTCGAATTTGGCTCTTTTACC 60
DB |||||
34 ATGGACCGTTTCGATCGCATGAGGCGCTTTGCCGCTGTGTGATACCGGACGCTTCACC 93
QY 61 GCGCGCCGACAGCTACAGTACAGCTGTTGTCATCAGTACAGCTATCAAACTG 120
DB |||||
94 AAGCGCGGAGACGCTGACATGAGCGCCGACGCGTCAACCGGCAAGCTCACTGACCGCC 153
QY 121 GAAGATGAGTTGAGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTGACCGAA 180
DB |||||
154 GAAGCGCGCTGCGGTAGCGCTGCTCAACCGCACACCGGCAAGCTCACTGACCGCC 213
QY 181 GCGGTAGATTTACTACAGGCTGCGCTGATGCTTTCATGAAGTACGAGTTCAT 240
DB |||||
214 GACGCGCGCGCTACTACAGCGGGTGTACGCTGCTGGCGGAATGACGATCGCGAG 273
QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGACGCTACGATGCTGTTCTTCA 300
DB |||||
274 ACCAGCTCTCCAGCGCCCGCTGCGCGCCCGCGGCGTCTCCGATGACGTACCGAGC 333
QY 301 ACTATGGCAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGAAATACCCAGGT 360
DB |||||
334 CCGCTGGCGGGTGTGCTGATCCGCGGTTGCGCGAGTTCATCGCGCTATCGGAG 393
QY 361 TTGAGCGTCAATCTGTACCGAATTCAGCGCCCGGACCTGATTCGCGAGGCTTGAT 420
DB |||||
394 ATCCAGATGACCTGGCGGTGAGCGATCGGCTGCTGATAGCGCGGGTTCGCGAC 453
QY 421 GTGGTATCGCGGTGCGCGGTGACAGGATTCAGGCTGTTTCCGCGCTGCTGGCGCG 480
DB |||||
454 TCGCTATCCGCGGCGGACCTGCGGACAGTCTGATAGCGCGGGTTCGCGAC 513
QY 481 ATGCCAATGGTGTGCGCGGAAAGCTATCTCACAAATACGGATACCGGAAAAA 540
DB |||||
514 CTGCACTGGGGTCTACCGGCAACCGGCTACTTTCAGCGCGCGCGGGTTCGCGAGCCAC 573
QY 541 CCGCGCGATTG 552
DB |||||
574 CCGCGGAACTG 585

RESULT 11

ABD05418

ID ABD05418 standard; DNA; 1047 BP.

XX

AC

XX ABD05418;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polynucleotide #4022.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.

OS Pseudomonas aeruginosa.

PN US5551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR P-PSDB; ABO71847.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 4022; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biotech technology. Sequences ABD01397-
XX ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

SQ Sequence 1047 BP; 167 A; 361 C; 355 G; 164 T; 0 U; 0 Other;

Query Match 11.9%; Score 110.4; DB 11; Length 1047;

Best Local Similarity 50.0%; Pred. No. 3.1e-23;

Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCTATGCGTGTGTCGCAAGTAGTTCGAATTTGGCTCTTTTACC 60

DB |||||
44 ATGGAACGCTTCGATCGCATGAGGCGCTTTGCCGCTGTGTGATACCGGACGCTTCACC 103

QY 61 GCGCGCCGACAGCTACAGTACAGCTGAGCGTTTCGTTCATCAGTACAGCGGATCAAACTG 120

DB |||||
104 AAGCGCGGAGAGCGCTGACATGAGCGGACCGGACCGGCTACCCAGTTGGTCCAGCACTG 163

QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACAGGACGATTCGCTGACCGAA 180

DB |||||
164 GAAGCGCGCTGCGGGTACCGCTGCTCAACCGCACCGGCAAGCTCAACCTGACCGCC 223

QY 181 GCGGTAGATTTACTACAGGCGTCCGCTGATGCTTCATGAAGTACGAGTTCAT 240

DB |||||
224 GACGCGCGCGCTACTACGAGCGGGTGTACGCTGCTGCGGAACTGAGCGATGCGGAG 283

QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGACGCTACGCACTGCTGCTTCTTCA 300

Db 284 ACCAGGCTCTCCAGCGCCGCTTGGCGCCGCGGCGTCTCCGATCGACGTACCGAGC 343
Qy 301 ACTATGGCACAATATGTTCTCGCGGGCTGACAGCCAAATCTGAAAGAAATACCCAGT 360
Db 344 CCCTGGCGGGTGTCTGATCCCGGGTTCGCGAGTTCCATGCGCGCTATCCGAG 403
Qy 361 TTGAGCGTCAATCTGTTTACCGGAATCCAGCCCGCGACCTGATCCGACGGTCTGGAT 420
Db 404 ATCCAGATCGACCTGGCGGTGAGCGATCGCTCGTACCTGATCGACGAGAATCGTCAC 463
Qy 421 GTGGTATCCGCTCGCGGGTTCAGGATCCAGCCTGTTTTCGCGCGTCTGGCGCG 480
Db 464 TGGGTATCCGCGCGCGGAACCTGCGCACCTGCTGATAGCGCGGGGTGCGCGAC 523
Qy 481 ATCCCAATGTGTGTGCGCGCGGAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
Db 524 CTGCAACTGGGGTCTACGCGGCACCGCTACTTGCAGCGCGCGCGGTGCGAGCCAC 583
Qy 541 CCGCGCGATTG 552
Db 584 CCGCGCGAACTG 595

RESULT 12

ABD05289/c

ID ABD05289 standard; DNA; 1089 BP.

XX AC ABD05289;

XX AC

XX DT 29-JUL-2004 (first entry)

XX XX

XX DE Pseudomonas aeruginosa polynucleotide #3893.

XX XX

XX KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;

XX KW antibacterial.

XX XX

XX OS Pseudomonas aeruginosa.

XX XX

XX PN US651795-B1.

XX XX

XX PD 22-APR-2003.

XX XX

XX PF 18-FEB-1999; 99US-00252991.

XX XX

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX XX

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX XX

XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX XX

XX DR WPI; 2003-615309/58.

XX DR P-PSDB; ABO71718.

XX XX

XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX PS Disclosure; SEQ ID NO 3893; 455pp; English.

XX XX

XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using biochip technology. Sequences ABD01397-

CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 1089 BP; 168 A; 372 C; 372 G; 177 T; 0 U; 0 Other;

Query Match 11.9%; Score 110.4; DB 11; Length 1089;
Best Local Similarity 50.0%; Pred. No. 3.1e-23;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

Qy 1 ATGGAACGACTAAACCGCATGTCGGTGTGTTTCCAAAGTAGTTGAATTGGCTCTTTTACC 60
Db 981 ATGACCGTTTCGATGCGATGAGGCTTTCGCGTGTGTTGATACCGGACGTTTACC 922
Qy 61 GCCGCCCCAGACAGCTACAGATGAGGTTTCTCCATCAGTCAGAGGTTTCAAACTG 120
Db 921 AAGCGCGGAGAGCGTGCACATGAGCGCGCACCGGTCACCCAGTTGGTCCAGCAACTG 862
Qy 121 GAGATGAGTTGAGGTAAAGCTGTTAAACGCTAGCACACGACGATTTGCCCTGACCGAA 180
Db 861 GAGCGCGGCTGCGGTTACGCTTCTCAACCGCACCGGCAAGCTCAACCTGACCGCC 802
Qy 181 GCCGTAGAATTTACTACAGGGCTGCCGTCTGATGCTTCAAGTGCAGGATGTTTCA 240
Db 801 GACGGCGCGCTACTACGAGCGGTGTTACGTTCTGTGCGGAACTGGACGATGCGGAG 742
Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTACGATTTGGCTTCTTCA 300
Db 741 ACCAGCCTCTCCAGCGCGCCCTCGGCGCCCGCGGGCTCTCCGATCGACGTACGAGC 682
Qy 301 ACTATGACACAAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGT 360
Db 681 CCGCTGCGCGGCTGCTGCTGATCCCGGCTTCCCGAGTTCCATGCGCGCTATCCGAG 622
Qy 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCGACCTGATTCGCGACGTTCTGGAT 420
Db 621 ATCCAGATCGACCTGCGGCTGAGCGATCGGCTCGTCCGCTGATCGACGAGAACGTCGAC 562
Qy 421 GTGGTATCCGCGCGCGGTTGAGGATTCAGGCTTCAGGCTTTTCCGCGCGCTCTGGGCGG 480
Db 561 TGGTCACTCGCGCGGGAACCTGCGGACCACTGCTGATAGCGCGGGTTCGCGAC 502
Qy 481 ATGCAATGTGTGTGCGCGGAAAGTATCTTCAACAATACGATACCGGATAACCGAAAA 540
Db 501 CTGCAACTGGGGTCTACGCGGACCGGCGCTTCTGAGCGCGCGCGGCTGCGGAGCCAC 442
Qy 541 CCGCGCGATTG 552
Db 441 CCGCGCGAACTG 430

RESULT 13

ACA44271

ID ACA44271 standard; DNA; 945 BP.

XX AC ACA44271;

XX XX

XX DT 19-JUN-2003 (first entry)

XX XX

XX DE Prokaryotic essential gene #25928.

XX XX

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;

XX KW drug design; gene.

XX XX

XX OS Pseudomonas putida.

XX XX

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX XX

XX PF 21-MAR-2002; 2002WO-US009107.

XX XX

Matches 412; Conservative 0; Mismatches 459; Indels 8; Gaps 2;

QY 10 CTAACAGCATGTCGGTGTTCGCCAAGTAGTGAATTTGGCTCTTTACCGCCGCCGCC 69

Db 37 CTTACGGGTGGCCATTTTATCGCGGTGGTAGATGCGCGAGTTTACCTGTCGCCGA 96

QY 70 AGACAGTACAGATGAGCGTTTCCTCCATCAGTCAGACGGTATCAAACTGGAAGATGAG 129

Db 97 CGCGCGTGGGACAGACAGGCGGTGGTCACTTTAAGTCAGCGAGCTGGAAATGAG 156

QY 130 TTCCAGGTAAAGTGTAAACCGTAGCAGCAGCAATTTGGCTGACCGAAGCCGGTAGA 189

Db 157 CTGGGGTACGCTGCTGCTGCTTTCGACCGCGCGTTTTCGGCTCACTGACCGCGGGTG 216

QY 190 ATTACTACAGCGCTGCGCTGTATGCTTTCATGAAGTCAGGATGTTTCATCAGCACTG 249

Db 217 CTGTTCTATCAGCGCGGGTGGCTCTTGAACGCGCGGAAATCTCCAGATGAGTG 276

QY 250 TATGCTTCAATTAACACCCCATTCGGACGCTAGCAATTTGGTGTTCCTCAACTATGCA 309

Db 277 CGGCGCAGCATAGCGCTTAAGCGGGAGCTACGGATCACCAACGCGCGGATGATGCG 336

QY 310 CAAAGTGTTCGCGCGGTGACAGCCAAATGCTGAAGATACCCAGTTTGGAGCTG 369

Db 337 GCGAGGTATATTCAGCGCTGGCGCGGTTCGCCCTGCTCATCCGCGCTGCGGTG 396

QY 370 AATCTGTTACCGGAATTCAGCGCCCGACCTGATTCGCGAGCGTCTGGATTTGGTATC 429

Db 397 CGCACGTTCTCTCTCCCATCATGCGACCTCATTTTCAAGCAATTTGATGCGGATA 456

QY 430 CGGTCGCGCGGTTCAGAGATTCAGCTGTTTTCGCGCGTCTGGCGCGGATGCAATG 489

Db 457 CGCCTCGTACGCTGCGGATTCCTCCCTATCGCGGACGCGGATAGCCAGTTTCGCGATC 516

QY 490 GTGCTGCGCGCGAAGATCTATCTCACAAATAGGCATACCGGAAACCCGC--- 545

Db 517 CTGCGCTGCGCTCTCGGCTGGCTGGCAAGCCATCCGCTGAGACGCTGCTCCGATG 576

QY 546 ---CGATTGAGTAGTCAATTCATGCTTGAATACAGCGTGGCGCGCGCAATGAATTTGA 602

Db 577 GCGCAAGCTGAGTGGATTATTCAGAGCGCTGCGCGACGCGCTGCGCTGGCAGCTGCG 636

QY 603 ACTGATCGCACCGAGGATCTGATCGCTGATCCCAAGGAGATTTGTGACTAA 662

Db 637 ACCGATCACAGACAGGTGATTTGCGATAGCCAGCGCGCGGTTTCTG-CTGA 695

QY 663 TGATCCGATGACGCTGGTGGCTGCTGAGCGGGTGGCGGATGCGCTAGCTGCGCGT 722

Db 696 TAGCGCTACGCGCTGATGAGCTTTGCTCTGGCGGATGCGCGCTGCTGCTGCGCGC 755

QY 723 GATGTGGGTATCAACAGATCAATCTGGGGAGCTGGAGATCTCTGCTGCGCGTTTACA 782

Db 756 ATGGCTGTAGCAAAAGAGTGCGGAGCGGGAGCTGGTGGCGCTCTCTGCGGAAATACA 815

QY 783 GTGAGATCCACGCGCGGTATTCGTTATATATACCGAAAGATTAAGCTGCGCTGAAAGT 842

Db 816 TTTCCACACAGCGGGGTGATGCTCTTATCCGATCGCAGCATTTGCGCGCGCGGT 875

QY 843 ACAGGTGCTGATCACTGCTGACGATTTATTTGTTGA 881

Db 876 GCGGGCATTTATTGACTTCTCTGCGCGAGAGGTGCGGTTA 914

RESULT 15

ABD02170/c

ID ABD02170 standard; DNA; 912 BP.

XX AC

XX AC

XX AC

XX 29-JUL-2004 (first entry)

XX DE

XX Pseudomonas aeruginosa polynucleotide #774.

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
 KW antibacterial.
 XX Pseudomonas aeruginosa.
 OS US6551795-B1.
 PN 22-APR-2003.
 XX 18-FEB-1999; 99US-00252991.
 XX 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 PI WPI; 2003-615309/58.
 DR P-PSDB; ABO68599.
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX Disclosure; SEQ ID NO 774; 455pp; English.
 PS The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABD01397-
 CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX Sequence 912 BP; 164 A; 296 C; 307 G; 145 T; 0 U; 0 Other;

Query Match 11.6%; Score 108; DB 11; Length 912;
 Best Local Similarity 47.3%; Pred. No. 1.6e-22;
 Matches 361; Conservative 0; Mismatches 400; Indels 3; Gaps 1;

QY 1 ATGGAACGACATAAAACGATGTCGGTGTTCGCCAAGTAGTGAATTTGGCTCTTTACC 60

Db 768 ATGATTCGCTGAAGCATGGCGATCTTCGCCATCTGTGTGACAGGGCTCGATGGCA 709

QY 61 GCCGCCGCCAGACAGCTACAGATGAGGTTTGTTCATCATAGTCAGACGGTATCAAACTG 120

Db 708 GCCGCCGCCAGAGCTTGGCATGACCCCTTCGCGAGTCAAGCCAGCAGATCCGAAAGCTG 649

QY 121 GAAGATGAGTTGACGATTAAGCTGTTAAACCGTAGCACAGCAGCATTTGCCCTGACCGAA 180

Db 648 GAGAGCGCGCCAGGTCATTTGCTGATCGCACACCCCGCGGCTGACCTTACCGAG 589

QY 181 GCCCGTGAATTTACTACAGGGCTGCCGTGATGCTTTCATGATGACGAGATGTTTAT 240

Db 588 GCCCGAGGCGTCTTATCGAGTTTGGCGAGATGCTGCGGATCGCCGAGGAGCGGAG 529

QY 241 GAGCACTGTATGCTTCAATACACCCCATCGGAGCTAGCATTTGGCTTTCTTCA 300

Db 528 CGCGCGCTTGGCGAATGGCGGATGCGCGGTGGGTGAGTTGCGCTTGGCGCGCGGCTG 469

QY 301 ACTATGCAACAAATGTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360

Db 468 GGGTCTCCGCGAGCTGATCACCCAGGCGCTGAAACCGTTGCTGGAGAACCCAGCGCAG 409

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 15:34:11 ; Search time 104.036 Seconds
(without alignments)
6353.876 Million cell updates/sec

Title: US-10-759-889-1

Perfect score: 930
Sequence: 1 atggacactaaacgcacat.....ggcgagggaagaagaagtaa 930

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	621.4	66.8	933	4	US-09-489-039A-6503
2	122	13.1	948	4	US-09-252-991A-643
3	121.8	13.1	1026	4	US-09-252-991A-5186
4	112.6	12.1	927	4	US-09-489-039A-2078
5	110.4	11.9	957	4	US-09-252-991A-3952
6	110.4	11.9	1047	4	US-09-252-991A-4022
7	110.4	11.9	1089	4	US-09-252-991A-3893
8	108.6	11.7	915	4	US-09-489-039A-2786
9	108	11.6	912	4	US-09-252-991A-774
10	107.6	11.6	1017	4	US-09-252-991A-11975
11	107.6	11.6	1104	4	US-09-252-991A-11906
12	106.4	11.4	1029	4	US-09-252-991A-13686
13	106.4	11.4	1071	4	US-09-252-991A-13459
14	106.4	11.4	1809	4	US-09-252-991A-13887
15	103.2	11.1	933	4	US-09-489-039A-454
16	102.8	11.1	777	4	US-09-252-991A-5250
17	102.8	11.0	591	4	US-09-134-000C-11
18	101.4	10.9	915	4	US-09-252-991A-14968
19	101.4	10.9	1011	4	US-09-252-991A-14782
20	100.4	10.8	1062	4	US-09-252-991A-9086
21	100.2	10.8	1008	4	US-09-489-039A-3346
22	98.6	10.6	792	4	US-09-252-991A-12012
23	95.4	10.3	906	4	US-09-252-991A-8825
24	95.2	10.2	966	4	US-09-252-991A-3656
25	95.2	10.2	1017	4	US-09-489-039A-1578
26	95.2	10.2	1026	4	US-09-252-991A-3611
27	95.2	10.2	1350	4	US-09-252-991A-3851

ALIGNMENTS

RESULT 1

US-09-489-039A-6503

; Sequence 6503, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Bregon et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004.001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 6503

; LENGTH: 933

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-6503

Query Match 66.8%; Score 621.4; DB 4; Length 933;

Best Local Similarity 79.8%; Pred. No. 1.7e-190;

Matches 733; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

Qy	1	ATGGAACGACTAAACCGCATGTCGGTGTTCGCAAAAGTAGTTGAATTTGGCTCTTTTACC	60
Db	4	ATGGAACGATTAAACCGCATGTCGGTGTTCGCAAAAGTAGTTGAATTTGGCTCTTTTACC	63
Qy	61	GCGCGCCCGCAGACGCTACAGATGAGCGTTTCCTCCATCAGTCAGACGGTATCAAACTG	120
Db	64	GCAGCGCCCGCAGCTTCAGATGAGCGTTTCCTCCATCAGTCAGACGGTATCAAACTG	123
Qy	121	GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACAGCAGCATTTGCGTACCGGAA	180
Db	124	GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACAGCAGCAGCATTTGCGTACCGGAA	183
Qy	181	GCGGTAGATTTACTACGAGGCTGCGTCTGATGTTTCATCAAGTGAGGATGTTTCAT	240
Db	184	GCGGTAGATTTACTACGAGGCTGCGTCTGATGTTTCATCAAGTGAGGATGTTTCAT	243
Qy	241	GAGCACTGTATGCTTTCAATACACCCCATCGGACGCTAGCGCATTTGCTGCTTTTCA	300
Db	244	GAGCACTGTATGCTTTCAATACACCCCATCGGACGCTAGCGCATTTGCTGCTTTTCA	303
Qy	301	ACTATGCGCAAAATGTTCTCGCGGGCTGACGCCAAATGCTGAAAGTAATCCAGGT	360
Db	304	ACTATGCGCAAAATGTTCTCGCGGGCTGACGCCAAATGCTGAAAGTAATCCAGGT	363
Qy	361	TTGAGCGTCAATCTGGTTACCGGAATTCACGCCCGGCTGATTCGCGACGCTCTGAT	420

Db 364 CTCCTCTGTTAAATCTGGTCAACGGCATACCTGCGCCAGACCTGATCGCCGACGGCCTTGAC 423
 Qy 421 GTGGTGATCCGGCGTGGCGGCTTGAGGATTCACAGCTGTTTCCCGCGCTGCGGCGC 480
 Db 424 GTGGTGATCCGGCGTGGCGGCTTGAGGATTCACAGCTGTTTCCCGCGCTGCGGCTCA 483
 Qy 481 ATGCCAATGGTGTGTCGCGCCGCGAAAGCTATCTCACAAATACGGCATACCGGAAAA 540
 Db 484 ATGCCAATGGTATATGTCGCGCCGCTAAAGCTATCTGCGCGAGCTGGCAATCCAGAGAA 543
 Qy 541 CCGCGCGATTTGAGTAGTCAATTCATGGCTTGATACAGCGTGGCGCGCGCAAGTAATTT 600
 Db 544 CCGCGCGACCTTGC CGGACATCGCTGGCTGGGAATATAGCTGCGCGCGGCAATGAGTTT 603
 Qy 601 GAACCTGATCGCACCGGAGGATCTCGACTCGCCTGATCCCAAGGAGATTTGTGACT 660
 Db 604 GTCAATTCGCGCGGAGGATCTCGACCCGCTGACCCCGCAGCGGCTTCTGACC 663
 Qy 661 AATGATCCGATGACGCTGTGCTGGCTGAGCGGCGGTGCGGATCGCCTACGTCGG 720
 Db 664 AACGACCTATGACGCTGTGCTGGCTGAGCGCGCGGTGGGTATCGCCTATGTACCC 723
 Qy 721 CTGATCTGGGTGATCAACAGAGATCAATCGTGGGAGCTGGAGATCTGCTGCGCGGTTAC 780
 Db 724 TTAATGTGGGATCGAAGAGATCAACCGCGTGACTTGATCTGCTGCCAGCTAT 783
 Qy 781 CAGTCAGATCCAGCGCGCTTTATGGTTATATACGAAAAAGATPAAGCTGCGCTGAAG 840
 Db 784 CAGTCAGATCCAGCGCGCTTACGCGCTGTATACGAAAAAGATPAAGCTGCGCTGAAG 843
 Qy 841 GTACAGCTGCTGATCAACTCGTGAAGGATTTTGTGAGTGGTGAATTTGTTTCAG 900
 Db 844 GTGAGGCTGTATCAACTATCTGACGGAGTATTTTGTGACGTGGGAGATCTATCAG 903
 Qy 901 GAGATGACGGCGCGGGA 919
 Db 904 GGCATCACGGCGCGGGA 922

RESULT 2

US-09-252-991A-643
 ; Sequence 643, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 643
 ; LENGTH: 948
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-643

Query Match 13.1%; Score 122; DB 4; Length 948;
 Best Local Similarity 46.9%; Pred. No. 4e-29;
 Matches 417; Conservative 0; Mismatches 470; Indels 3; Gaps 1;
 Qy 1 ATGCRACGCTAAACGCGATGCTGGTGTTCGCAAGTAGTGAATTTGGCTCTTTTACC 60
 Db 40 ATGATTCGCTGAAGGCGATGCGATCTTCGCCACTGTGGTGGCAAGGCTTCGATGCA 99
 Qy 61 GCGCGCGCCAGACGATACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 120
 Db 100 GCGCGCGCCAGACGCTGGGCGATGACCCCTTCGGCAGTCAGCCAGATCGCAAGCTG 159

Qy 121 GAAGATGAGTTGAGTAAAGCTGTTTAAACCGTAGCACACGAGCATTTGGCCTTGACCGAA 180
 Db 160 GAGAGCGCGCCAGGTCATTTGCTGTCATCGCACCCACCGCGGCTGACCTTGCACCGAG 219
 Qy 181 GCGGTGAGATTTACTACAGGGCTGCGCTGCTATGCTTTCATGAGTGCAGGATGTTTCA 240
 Db 220 GCGCGGAGCGGCTTCTATCGAGTTGCGCGAGATCTGCGCATGCGCGAGAGCCGAG 279
 Qy 241 GAGCAATCTATGCTTCAATTAACACCCCATCGGACGCTACGCAATTCGCTTCTTCA 300
 Db 280 CCGCGGCTTGGCGAATGGCGGATGCGCGGTGGTGGTGGCTTGGCGCTGCGCGCGGCTG 339
 Qy 301 ACTATGGCAAAATGTTCTCGCGGCTGACAGCCAAATGCTGAAGAAATACCCAGT 360
 Db 340 GGTTTCTCGCGCAGCTGATCACCCAGCGCTGAACCCGTTGTTGGAGAACCCGCCAG 399
 Qy 361 TTGAGGCTCAATCTGGTTACCGGAATTCAGCGCCGCACTGATTCGCCGACGCTCTGGAT 420
 Db 400 TTGCGCTGAGCTGTTCTTCAGGACGAGCGCATGATCTGCTGCGCGGCGATCGAC 459
 Qy 421 GTGGTGATCGCGCTGCGCGGCTTTCAGGATTCAGAGCTGTTTCCGCGCTGCGGCGC 480
 Db 460 CTGCGCATACGCTGCGCAATCTCGCGGCTGCGCGCTGCGCGCGCCACCTCGCGCAT 519
 Qy 481 ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACAAATACGCGCATACCGGAAAA 540
 Db 520 TGGAGCAGCTGCTTTCGCGCGCGCGCTTATTTGCGCCACGCGGCGCCATCAACGT 579
 Qy 541 CCGCGCGATTTGAGTAGTCAATTCATGCTT---GAATACAGCGTGGCGCGCCGCAATGAA 597
 Db 580 CCGGAGCAGTTGACCGAGCTGCACTGATTTCCCTGAACACCTCGAACACCCCTCAACCAC 639
 Qy 598 TTTGAATGATCCAGCGGAAAGGATCTGACCTGCTGATCCCAAGGAAAGTTTGTG 657
 Db 640 CTGACGCTCAGCGCGCGCGCGGGAAGTCTGCAAACTGCGCTGAGGACCGGTGGC 699
 Qy 658 ACTAATGATCCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
 Db 700 GCCAACGGGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
 Qy 718 CCGCTGATGCTGCTGATCAACGAGATCAATCTGCGGAGCTGGAGATCTGCTGCGCGCT 777
 Db 760 CCGCTGCGGAGTGGCGCGCGCTGCAACGAGCGGCTGCAACGAGCGGCTGCAACGAGCT 819
 Qy 778 TACGAGTCAGATCCAGCGCGGCTTATGCTTATATACCGAAAAAGATTAAGCTGCGGCTG 837
 Db 820 TGGAGGATTCGCGGCTCGGCATCTATCGGTGACGCCACGCGCGGAGCGCCAGCGGCC 879
 Qy 838 AAGTACAGTCTGATCAACTCGCTGACGAGTATTTTGTGAGTCTGG 887
 Db 880 AAGTGAAGTTCGCTGAGCGGCTGCGCGGAGCTGCGCGGAGCTTCCGCCAGCGG 929

RESULT 3

US-09-252-991A-5186
 ; Sequence 5186, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5186
 ; LENGTH: 1026
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-5186

```
Query Match      13.1%; Score 121.8; DB 4; Length 1026;
Best Local Similarity 47.1%; Pred. No. 4.9e-29;
Matches 409; Conservative 0; Mismatches 457; Indels 3; Gaps 1;

QY 1 ATGGACGACTAAACGCGATGCGGTGTTTGGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 106 ATGGACGCGCTACCGCGCTCGCTTTTCGTGGAAGTCATCGACGCGGACGACGACC 165
QY 61 GCGCGCGCGACGAGCTACAGATAGCGTTTCGTCCATCAGTCAGACGATATCAAACTG 120
DB 166 GCGCGCGCGAGCGTCTGACATGTCGCGGGGATGTTCTCGCGTACCTCGCGGAGTTG 225
QY 121 GAAGATGAGTGCAGGTAAAGCTGTTAAACCTAGCACACGAGCATTTGCCGTACCGAA 180
DB 226 GAGGAGTGGTTCGGCGACGCTGTCACCGGACGACCGCGCGCTGAGCCTGACCGAC 285
QY 181 GCGGTAGAAATTTACTACAGGCGTGCCTGCTATGCTTTCAAGATGAGGATGTCAT 240
DB 286 GCGCGCGCGAGCTGTTGCGCGAGTGCCTGCGAGATGCTCGCGTGGCGGACCGATGAG 345
QY 241 GAGCACTGTATGCTTTCAATACACCCCATCGGACGCTACGATTTGGCTTTCTTCA 300
DB 346 GCGCGCGCGACGCGCGCGACAGCGCGCGGCACTCTCGCATCACACGAGCGCTG 405
QY 301 ACTATGGCACAATAATTTCTCGCGGCGTACAGCCAAATGCTGAAAGATACCCAGT 360
DB 406 TCTTTCGCCAGCGCTGCTGACCGCTGCGCGCGCGCGCTTCTGCGGGCTACCGGCG 465
QY 361 TTGAGGCTCAATGCTTTACCGGAATTCAGCGCCCGACCTGATTTGCGAGCGTTCGAT 420
DB 466 ACCGCCATCGATCGAGTCAACAGCCAGGTGCTCAACCTGGTGGAGGAGCGCATCGAC 525
QY 421 GTGGTATCGGCTGCGCGCTTGAGGATTCAGCGTGTTCAGCGTGTTCGCGCGCGCG 480
DB 526 CTAGCCCTCGGATCGCCAAACAGCTCGACCCCACTGATCGCGCGCGCGCTCGCGAA 585
QY 481 ATGCCAATCGTGTGCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAAAA 540
DB 586 TGCCTGCTGATGATCGCGCGCGCGGAGTACCTGCGCGCGCGCACCGCGCGCGCG 645
QY 541 CCGCGGATTTGAGTAGTATTCATGCTTGAATACAGCGTGGCGCGCGCGCGAGTAA 600
DB 646 CCGGAGGACTCGGCTGCGACAACTGCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCT 702
QY 601 GAAGTATCGCGCGGAGGAGTCTGACTCGCTGCTGATCCCAAGGAGATTTGTGACT 660
DB 703 TGGCAGTTTCAGCGCGGACGCGCGGACGAGCTGCGGCTGGGTGGGAGTCTAGCGCG 762
QY 661 AATGATCCGATGACGCTGTGCTGCTGCTGAGCGCGGCTGCGGAGTCCCTTACGTCG 720
DB 763 AACGAATCCACCGTGTGCTGCTGAGGCGCGCGCGCGGCTGCGGAGTACAGCAGCGCG 822
QY 721 CTGATGCTGCTGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGCTGCGGTTAC 780
DB 823 CTGACTCGGAGCGCGCTGATTCGTAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTG 882
QY 781 CAGTCAGATCCAGCGCGGTTTATGCGTTATATACCGAAAAAGATAAGTTCGCTGAAG 840
DB 883 TCGCGCGAGTGTCTGGCATCCACGCGGTGATCGGCTCGCGCGCGAGATGCGCGCGCG 942
QY 841 GTACAGTGTGATCAACTCGCTGACGGA 869
DB 943 CTGCGGCGCTGCTGACTTCTGTGTGCA 971
```

RESULT 4

US-09-489-039A-2078
; Sequence 2078; Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 2078  
; LENGTH: 927  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-2078
```

```
Query Match      12.1%; Score 112.6; DB 4; Length 927;
Best Local Similarity 45.7%; Pred. No. 4.3e-26;
Matches 394; Conservative 0; Mismatches 469; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCGATGCGGTGTTTTCGCAAACTAGTTGAATTTGGCTCTTTTACC 60
DB 16 ATGGAACAATTTACGAGGATGGAGACGTTTATTGCCGTGGTGGAGTCGCGCAGCTTACC 75
QY 61 GCGCGCGCGACAGCTACAGATAGAGCGTTTCGTTCATCAGTCAGACGCTATCAAACTG 120
DB 76 GCGCGCGCGTGCAGCGCTGGGCGTGTGCGCGGTGATGTTGGGAATATATCGCCACGCTG 135
QY 121 GAAGATCAGTTGCAAGTAAAGCTGTTAAACGCTAGCACACGAGCATTTGCCCTGACCGAA 180
DB 136 GAAGACGAGTGGCGAGCGCTGCTGAGCGCTTAACCCGACGCGAGAGCTCACCAGC 195
QY 181 GCGGTAGAAATTTACTACAGGCGTGCCTGCTGATGCTTCATGAAGTGCAGGATGTTTAT 240
DB 196 GCGCGCGCGTCTATTTTGAAGGCGAAACGCGGTGATGAGCAGGTCTCCATCGCCGAA 255
QY 241 GAGCAACTGTATGCTTTCAATACACCCCATCGGAGCGTACGCTGCTGCTGCTGCTGCT 300
DB 256 AGCGCGTTCAGCGCGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 315
QY 301 ACTATGCGCACAAATGTTCTCGCGCGCGCTGACAGCCAAATGCTGAAAGATATCCACG 360
DB 316 TCGTTTCGCGCGCGAGCTGATCGCGCGCGCTGACCGCGCACTTCTCGAGGCGTGGCTG 375
QY 361 TTGAGCTGCAATCTGCTTACGGAAATTCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 376 GTCGCGGTGGAATCTGATCTGATTAACCGGATGCTGATCTGCTGCTGCTGCTGCTGCT 435
QY 421 GTGCTGATTCGCGTTCGCGCGGTTCAGGATTCAGAGCTGTTTCCGCGCGCTCTGGGCG 480
DB 436 CTGCGGATTCGATTCGCGCGAGATCCATCAGAGGATCTGCTGCTGCTGCTGCTGCTG 495
QY 481 ATGCCAATGCTGTGTCGCGCGGAAAGCTATCTCACAAATACCGCATACCGGAAAAA 540
DB 496 TATCGGATGCTGATCTGCGCGCGCGCGCTTATCTGCGCGGTACGCGACCGCGCGAG 555
QY 541 CCGCGGATTTGAGTAGTATTCATGCTTGAATACAGCGTGGCGCGCGCGCGCGCGCG 600
DB 556 CCAGAGGATCTCGCGGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
QY 601 GAATGATTCGACCGGAGGAGTCTGACTCGCTGCTGATCCCAAGGAGATTTGTGACT 660
DB 616 TGGCGCTGCGCGCGGTAGAGGAGAGTGCCTGGAACGCGATGCGGTTTTTGGCGTGC 675
QY 661 AATGATCCGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 676 AACGATGGCTACGCGCTGCGCGCGCGGATCGCGCGCGCGCGCGCGCGCGCGCGCG 735
QY 721 CTGATGCTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCCTGCTGCGCGGTTAC 780
DB 736 GAGTACTGCTGCGCGATGCTGCGCGAGCGGAGCTGTTGCTGCGGTGCTGGAGGCTGG 795
QY 781 CAGTCAGATCCAGCGCGGTTTATGCGTTATATACCGAAAAAGATAAGTTCGCTGAAG 840
DB 796 ACTCGGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 855
```

QY 841 GTACAGTCTGATCACTCGCT 863
Db 856 CTGACGCACTTTATCGCCACCT 878

RESULT 5

US-09-252-991A-3952
; Sequence 3952, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3952
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3952

Query Match 11.9%; Score 110.4; DB 4; Length 957;
Best Local Similarity 50.0%; Pred. No. 2.3e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACCGCATGTCGGTGTGTTGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 34 ATGACCGGTTTCGATCGGATGACGGCTTTGCCGTGTGTGATACCGGAGCTTACC 93
QY 61 GCGCGCCGACAGCAGTACAGATGAGCGTTTGTCTCATCAGTCAGACGATATCAAACTG 120
Db 94 AAGCGCGCGAGACGCTGCACATGACGACCGCAGCTCACCAGTTGGTCCAGCAACTG 153
QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTTGACCGAA 180
Db 154 GAAGCGCGCTCGGGTACGCTGCTCAACCGCACACCGGACGCTCAACTGACCGCC 213
QY 181 GCGGTAGAAATTTACTACAGGCTCGCGTGTATGCTTTCTATGAAGTGAGGATGTTTCA 240
Db 214 GACGCGCGCTTACTACGAGCGGTGTGACGTTCTGCTGGCGGAACCTGGACGATCCGAG 273
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTTCTTCA 300
Db 274 ACAGCGCTCTCAGCGCGCCCTTGGCGCCCGCGGCGTCTCCGATCAGCTACCGAGC 333
QY 301 ACTATGGCACAAATTTCTCGCGGGCTGACAGCCAAATCTGAAAGATACCCAGT 360
Db 334 CGCTGGCGGCTGCTGTGATCCCGGGTTGCCGAGTTCCATGCGGCTATCCGGAG 393
QY 361 TTGAGCGTCAATTTGGTTACCGGAATTCAGCCCGGACGCTATTCGCGACGCTGCGAT 420
Db 394 ATCCAGATCGAATCGGCGGTGACGATCGGCTCGTGCACCTGATCGACGAGAACGTCAC 453
QY 421 GTGGTATCGCGCTCGGCGGTGTGACAGGATTCAGGCTGTTTCCCGCGCTCTGGCGCG 480
Db 454 TGGCTCATCGCGCGCGGAACTGCGCGACCGCTGATAGCGCGCGGCTCGCGAC 513
QY 481 ATGCCAATGTTGTGCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAATA 540
Db 514 CTGCAACTGGGGTCTACCGCGGACCGGCTTACTTGCAGCGCGCGGCTGCGGAGCCAC 573
QY 541 CCGCGCGAATTG 552
Db 574 CCGCGGAACTG 585

RESULT 7

US-09-252-991A-3893/c
; Sequence 3893, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

RESULT 6

US-09-252-991A-4022
; Sequence 4022, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4022
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4022

Query Match 11.9%; Score 110.4; DB 4; Length 1047;
Best Local Similarity 50.0%; Pred. No. 2.4e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACCGCATGTCGGTGTGTTGCCAAAGTAGTTGAATTTGGCTCTTTTACC 60
Db 44 ATGACCGGTTTCGATCGGATGACGGCTTTGCCGTGTGTGATACCGGAGCTTACC 103
QY 61 GCGCGCCGACAGCAGTACAGATGAGCGTTTGTCTCATCAGTCAGACGATATCAAACTG 120
Db 104 AAGCGCGCGAGACGCTGCACATGACGCGCAGCGTCCACCGATTTGGTCCAGCAACTG 163
QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTTGACCGAA 180
Db 164 GAAGCGCGCTCGGGTACGCTGCTCAACCGCACACCGGACGCTCAACTGACCGCC 223
QY 181 GCGGTAGAAATTTACTACAGGCTCGCGTGTATGCTTTCTATGAAGTGAGGATGTTTCA 240
Db 224 GACGCGCGCTTACTACGAGCGGTGTGACGTTCTGCTGGCGGAACCTGGACGATCCGAG 283
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTTCTTCA 300
Db 284 ACAGCGCTCTCAGCGCGCCCTTGGCGCCCGCGGCGTCTCCGATCAGCTACCGAGC 343
QY 301 ACTATGGCACAAATTTCTCGCGGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360
Db 344 CGCTGGCGGCTGCTGTGATCCCGGGTTGCCGAGTTCCATGCGGCTATCCGGAG 403
QY 361 TTGAGCGTCAATTTGGTTACCGGAATTCAGGCTGTTTCCCGCGCTCTGGCGCG 420
Db 404 ATCCAGATCGAATCGGCGGTGACGATCGGCTCGTGCACCTGATCGACGAGACGTCGAC 463
QY 421 GTGGTATCGCGCTCGGCGGTGTGACAGGATTCAGGCTGTTTCCCGCGCTCTGGCGCG 480
Db 464 TGGCTCATCGCGCGCGGAACTGCGCGACCGCTGATAGCGCGGCTCGCGAC 523
QY 481 ATGCCAATGTTGTGCGCGCGGAAAGCTATCTCACAAATACCGCATACCGGAAATA 540
Db 524 CTGCAACTGGGGTCTACCGCGGACCGGCTTACTTGCAGCGCGCGGCTGCGGAGCCAC 583
QY 541 CCGCGCGAATTG 552
Db 584 CCGCGGAACTG 595

```

: / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: /
: / TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: /
: / FILE REFERENCE: 107196.136
: /
: / CURRENT APPLICATION NUMBER: US/09/252,991A
: /
: / CURRENT FILING DATE: 1999-02-18
: /
: / PRIOR APPLICATION NUMBER: US 60/074,788
: /
: / PRIOR FILING DATE: 1998-02-18
: /
: / PRIOR APPLICATION NUMBER: US 60/094,190
: /
: / PRIOR FILING DATE: 1998-07-27
: /
: / NUMBER OF SEQ ID NOS: 33142
: /

```

; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 3893
 ; LENGTH: 1089
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3893

Query Match 11.9%; Score 110.4; DB 4; Length 1089;
Best Local Similarity 50.0%; Pred. No. 2.4e-25;
Matches 276; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

1	ATGGAACGACATAAAACGCGATGTCGGTGTGTTGGCAAAAGTAGTTGAATTTGGCTCTTTTACC	60
981	ATGGACCGTTTCGATCGGATCGCAGGCTTTGCCCGTGTGGTGATACCGCAGCTTCACC	922
61	CCCGCCGCACACAGCTACAGATGAGGTTTGGTCATCAGTCAGACGGTATCAAAATCG	120
921	AAGCGCGCGAGACGCTGCGACATGAGCGGCACAGCGTCACCCAGCTGGTCCAGCAACTG	862
121	GAAGATGAGTTGCGAGGTAAAGCTGTTAAACCGGTAGCACACGCGCATTTGGCTGACCGAA	180
861	GAAGCGCGGTGCGGGTACGCTTGCTCAACCGCACCGCGGAGAGCTCAACTGACGCC	802
181	CCCGGTAGAAATTACTACAGGCGTGCCTGCTATGCTTCATGAAGTCAGAGTGTTCAT	240
801	GACGGCGCGCTACTACGAGCGGTGGTAGCTCTGCTGGCGAACTGCGAGTGCAGG	742
241	GAGCAACTGATGCTTTCATATAACACCCCATCGGACGCTACGGATTGGCTGTCTTCA	300
741	ACGAGCTCTCCAGCGCGCGCTGGCGCGCCCGCGCGCTCCCGATCGAGTACCGGAGC	682
301	ACTATGGCACAAAATGTTCTCGCGCGGCTGACAGCCAAAATGCTGAAAGAAATACCGAGT	360
681	CCGCTGCGCGCGCTGCTGCTGATCCGCGGTGCCCGAGTTCATGCGCGCTATCCGGAG	622
361	TTGAGCTCAATCTGGTTACCGGAATTCAGGCCCGCAGCTGATTGCCGAGGTGTGAT	420
621	ATCCAGATCGACCTGGGCGGTGACGATCGGCTGCTCGACCTGATCGACGAGAAAGTCGAC	562
421	GTGTGTATCCGCGTCGCGCGTTGACAGGATTCAGGCTGTTTTCCCGCGCGTCTGGCGCG	480
561	TGCGTCAATCCGCGCGCGGAACTGCGCGACCAAGTCGCTGATAGCGCGCGGCTCGGCGAC	502
481	ATGCCAATGGTGTGTCGCGCGGAAAGCTATCTCACACAATAGGATACCGGAAAAA	540
501	CTGCAACTGGGGTCTACGCGGCACCGGCGTACTTGCAGCGCGCGGCTGCCAGGCCAC	442
541	CCCGCCGATTG	552
441	CCGCGGAACTG	430

RESULT 8

US 03-489-039A-2786
; Sequence 2786, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

```

; PRIOR FILING DATE: 1999-01-29
; ; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2786
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-483-039A-2786

```

```
Query Match      11.7%;      Score 108.6;   DB 4;      Length 915;
Best Local Similarity 46.9%;
Pred. No. 8.5e-25;
Matches 412; Conservative 0; Mismatches 459; Indels 8
```

10	QY	CTAATAACGCATGTCGGTGTGTTGCAAAAGTAGTTGAATTTTGGCTCTTTTACCGCCGCCGCCGCGC	69
37	Db	CTTCAGCCGCTGGGCATGTTTATCGCCGTGCTAGATCGCGCAGTTTACCCCTTGGCCGCGCA	96
70	QY	AGACAGCTACAGATGAGGGTTTCGTCCATCAGTCAGCGGTATCAAAAATGGAAGTGA	129
97	Db	GGCGCGCTGGGGCAGACCAAGCGCGTGGTCAGCTTTAAAGTACGGCAGCTGGAAAATGAG	156
130	QY	TTGCAGGTAAAGCTGTTTAAACCGGTAGCACACGAGCAATTGGCTGTGACCGAAACCCGGTAGA	189
157	Db	CTGGGGGTGACGCTGCTGCTGCGTTCGACGCGCGTTTGCCTGCTCACTGACGCCGGGGT	216
190	QY	ATTTACTACAGGGCTGCCGTGTAATGCTTCATGAAGTGCAGGATGTTTCATGAGCACTG	249
217	Db	CTGTCTTATCAGCGCGGGGTGGCGCTCTCTGAACGCGCGGAAAATCTCCAGATGAAGTG	276
250	QY	TATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTGGCTGTTCTTTCAACTATGGCA	309
277	Db	CGGGCCAGCCATACGGGCTTTAAGGGGGAGCTACGANTCACACACGCGGAGATATGCG	336
310	QY	CAAAATGTTTCGCGCGGGCTGACAGCCCAAAATGCTGAAAGATACCCAGTTTGTAGCGTTC	369
337	Db	CGCAGGTTATCATTTCCAGCGCTGGCGCGCTTCGCCCGTCTGTCATCCGCGCTGCGGGTG	396
370	QY	AATCTGGTTACCGGAAATTCAGCCCCGACCTCATTTGCGAGCGTCTGATCTGATGTTGATC	429
397	Db	CGCCAGCTTCTCTCTCCCATCATCGGACCTCATTTTCAAGCAGTTTGATGTGGCGATA	456
430	QY	CGCGTCGCGCGGTTTCAGGATTCAGCGCTGTTTTCGCCGCGTGTGGCGCGATGCCAATG	489
457	Db	CGCCTCGGTACGCTGGCGGATTCGCGTATCGGGCGACGCGGATACCCAGTTTCGCGATC	516
490	QY	GTGCTGTCGCGCGGAAAGGCTATCTCACAAATACGGCATACCGGAAAACCCGC	545
517	Db	CTGCGGCTCGGCTCTCGGGCTGGGTGGCAAGCCATCCGGTGACAGCGTGTCCGATCTG	576
546	QY	---CGATTTGAGTAGTCATTCATGGCTTGAATACAGCGTGCGGCGCCGCAATGAATTTGA	602
577	Db	CGCAAGCTGAGTGGATTATTCAGAGCGCGCTGCGACGCGCTGCGCTGGCGAGTGC	636
603	QY	ACTGATGCGACCGGAAGGATCTGACTGCTGATGCCATGCCACAGGAAGATTTGTGACTAA	662
637	Db	ACGGATCACCAGACAGAGTGGATTTTCGGGATAGCCAGCGCGCGGGTTTTCTG-CTGA	695
663	QY	TGATCCGATACGCTGGTGGCTGGCTGACGGCGGGTGC CGGGATCGCTACGTGCGCGT	722
696	Db	TAGCGCTACGGGCTGATGAGCTTTGCTCTGGCGGATCGGGCGTTCGCTTCTGCGCGC	755
723	QY	GATGCGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTTGCTGCCCGTTACCA	782
756	Db	ATGCTGCTGATACCAAAAGGTGGCGAGCGGGAGCTGGTGGCGCTCTGCGGGAATACCA	815
783	QY	GTCAAGTCCAGCCCGGTTTATCGGTTATATACCGAAAAGATAAGCTGCGCTGAAGT	842
816	Db	TTTCCACAGACGGCGTGATGCCCTTTATCCGACTCGAGCATTTTGCCGACCCGGT	875
843	QY	ACAGGTCGTGATCAACTCGCTACGGAATTTTGTGTA	881
876	Db	CGCGGCATTTATTGACTTCTCGCGCAGAGGTTCGGTTA	914

RESULT 9

US-09-252-991A-774/C
; Sequence 774, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 774
; LENGTH: 912
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-774

Query Match 11.6%; Score 108; DB 4; Length 912;
Best Local Similarity 47.3%; Pred. No. 1.3e-24;
Matches 361; Conservative 0; Mismatches 400; Indels 3; Gaps 1;
QY 1 ATGGAACGACTAAAGCGCATGTCGGTGTGTTGCCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 768 ATGGAATCGCTGAAGGCATGGCGATCTTCGCCACTGTGTGGACAAGGCTCGATGCA 709
QY 61 GCGCGCGCCAGACGCTACAGATGAGCGTTTGTGTCATCAGTCAGACGATCAAACTG 120
DB 708 GCGCGCGCCAGACGCTGGCGCATGACCTCTCGGCAGTCAGCCAGCATCCGCAAGCTG 649
QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCACGACGATTTGGCTGTACCGAA 180
DB 648 GAGAGCGCGCCAGCTGACTTTGTGTCATCGACACCGCGCGGTGACCTGACCGAG 589
QY 181 GCGGTAGAAATTTACTACAGGCTGCCGTGTATGTTCTATGAAGTCAGGATGTTTAT 240
DB 588 GCGCGCGAGCGCTTCTATCCAGTTGCGCGCATGCTGGCGATCGCCGAGGAAGCCGAG 529
QY 241 GAGCAACTGTATGCTTCAATACACCCCATCGGACGCTACGCAATTTGGCTGTCTTCA 300
DB 528 CGCGCGCTTGGCGAATGGCGCGATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 469
QY 301 ACTATGGCACAATGTTCTCGCGCGGCTGACAGCCAAAATGCTGAAGAATPACCCAGT 360
DB 468 GGGTTCTCGGCGACGCTGATCACCAGGCGCTGAACCCGTGCTGGAGAACCACCGCCAG 409
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGACCTGATTCGCGACCGTCTGGAT 420
DB 408 TTGCGCTGCGATGTTCTTCCAGGACGAGCGCATCGATCTGGTCCCGAGCGCATCGAC 349
QY 421 GTGGTATCGCGCTCGCGCGGTTGCAAGGATTCAGGCTGTTTTCGCGCGCTGCGCGCGG 480
DB 348 CTGGCGATAGCGTGGCGAATCTCGCGACTCGAGCTGTTGGCGCGCCACCTCGCGCAT 289
QY 481 ATGCAATAGTGTGTGCGCGCGAAGAGTATCTCACACATACGCGCATACCGGAATAA 540
DB 288 TGGAGCAGCGTGTCTGCGCGCGCGCGGCTATTTCGCGCAACGCGCGCCCATCAACCGT 229
QY 541 CCCGCGATTTGAGTAGTCAATTCATGCTTTGAAT---ACAGGTGGCGCGCCACATGAA 597
DB 228 CCGGAGCAGTTGACCGAGTGTGATGATTTCCCTGAAACACCTCGNACCATCTCAACAC 169
QY 598 TTGTAATGATCGACCGGAAGGATCTCGACTGCGCTGATCCCAAGGAAGATTTGTG 657
DB 168 CTGACGCTAGCGCGCGCGCGGGAAGTCTGCAAACTCGCGCTGAGGACCGGCTGGCC 109
QY 658 ACTAATGATCCGATGAGCTGTGGCTGGCTGAGCGCGCGGTGCGGGATCGCTACGCTG 717
DB 108 GCCAACGGGATGCTCGCGGTGGCTGAGTTCAACCTCGACGACTCGCGGTGCTTCTACCG 49

QY 718 CCGCTGATGGGTGATCAACGAGATCAATCGTGGGAGCTGGA 761
DB 48 CCGCTCCGAGAGTGGCGAGCGCTGAACGACGGCGCTGCA 5
RESULT 10
US-09-252-991A-11975
; Sequence 11975, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11975
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11975
Query Match 11.6%; Score 107.6; DB 4; Length 1017;
Best Local Similarity 46.1%; Pred. No. 1.9e-24;
Matches 399; Conservative 0; Mismatches 464; Indels 3; Gaps 1;
QY 2 TGGAACGACTAAACGCGATGTCGGTGTGTTGCCAAGTAGTTGAATTTGGCTCTTTTACC 61
DB 63 TGGACACATTTGCAAGGCATGCGGTGTTCGCCAGGTGTGGACACGCGCAGCTTCACTT 122
QY 62 CCGCGCGCACGACGCTACAGATGAGGTTCGTTCCTCATGTCAGACGCTATCAAACTGG 121
DB 123 CCGCGCGCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182
QY 122 AAGATGAGTTGCAAGTAAAGCTGTTAAACCGTAGCACACGACGATTCGCTGACCGAAG 181
DB 183 AGGCGCACCTGCGACGCGCTGCTGACCGCACCGCGCTGCGCGCTGCGCGCTGCGCGCTG 242
QY 182 CCGGTAGAAATTTACTACAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
DB 243 CCGCGCGCGCTTACTGGAACGCTGCGCGGATCTTCGAGGACATCGAGTTCGCGGAGG 302
QY 242 AGCAACTGTATGCTTCAATACACCCCATCGGAGCTACGCTACGCTGCTGCTGCTGCTG 301
DB 303 CCGAAGCGCGCGCGCGCACATCGCGCTGCGCGCTGCGCGCTGCGCTGCTGCTGCTGCTG 362
QY 302 CTATGCGACAATAATGTTCTCGCGGCTGACACCGCAAAATGCTGAAGAATACCCAGTT 361
DB 363 GCATGCGCGCGAGCATACCTGTCAGGCAACCGGACATGCTCGAGGAGGCGAGGACG 422
QY 362 TGACGCTCAATCTGCTTACCGGAATTCAGCGCCCGGACCTGATTTGCGCGAGCTGCTGATG 421
DB 423 TGGTGTGAGCTGACCTGTCAGGCAACCGGACATGCTCGAGGAGGCGAGGACG 482
QY 422 TGGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478
DB 483 TGCTGATCACCGCGAGCGCGAGTTGCGCGACTCGGAGTTCTGTCGCGCGAGCTGCTGCTG 542
QY 479 CGATGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538
DB 543 GCATTCACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
QY 539 AACCCGCGGATTTGAGTAGTCAATTCATGCTTGAATACAGCGTGGCGCGCGCAATGAAT 598
DB 603 CCGTGGACGACCTGGAGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
QY 599 TTGACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAAGGAAGATTTGTGA 658

Db 352 CGGACAGCTCTGACGAGATGAGGCGCTCGAGGAAACGCTGGCCCTCGCGCGGAGAGC 411
QY 268 CCATCGGACGCTAGCGATTGGCTGTTCTTTCAACTATGGACAAATATGTTCTGCCCGG 327
Db 412 CCGCGCGGCGCTGCGGATCAACGCGCTCGCCCTCATCTGACGCGGTGTGCGG 471
QY 328 CTGACAGCAAAATGCTGAAAGATACCAAGGTTGAGGGTCAATCTGTTACCGGAAT 387
Db 472 CACATCGGCGAGTTCGCGCGGTGTACCGGACATCCAGCTCGAATCAACACCAACGAC 531
QY 388 CCAGCCCCCGACTGATTCGCGACGCTCTGAGTGTGGTGTGATCCGGGTGCGGCGGTTCGAG 447
Db 532 CTGATCATGACCTGTTGAGCAGCAGCAGCTGGCGATCCGATCGGTGGCTCAGC 591
QY 448 GATTTCAGGCTGTTTTCGCGCGCTCTGCGGCGATGCCAATGGTGTGCGCGCGGAAA 507
Db 592 GACTCACCTCTGACGCGGCGCGGTGGGCGCGAGCGGCTGCACATCTCGCCAGCCCG 651
QY 508 AGCTATCTCACAAATACGCGCATACCGGAAACCGCGGATTTGAGTAGTCAATTCATGG 567
Db 652 GAGTACTGCGCGCAGCGGACGCGCGGAGCATGAGGGCTCGGAGGGGACACCCCTG 711
QY 568 CTGGAATACAGCGTGC 583
Db 712 CTCGGCTTACCAGC 727

RESULT 13

US-09-252-991A-13459/c
; Sequence 13459, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 13459

; LENGTH: 1071

TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13459

Query Match 11.4%; Score 106.4; DB 4; Length 1071;

Best Local Similarity 49.5%; Pred. No. 4.7e-24;

Matches 275; Conservative 0; Mismatches 281; Indels 0; Gaps 0;

QY 28 TTGCGCAAGTAGTGAATTTGCTCTTTTACCGCGCGCGCGGACAGTACAGATGAGC 87
Db 1035 TTGCTCAGTGTAGTCACACCGGTTCCATCAGCGCGCGCGGAAACAGATGGAGCTGACC 976
QY 88 GTTTCGTCATCAGTCAGACGGTATCAAAATCGGAAGATGAGTTGACGTTAAAGCTGTTA 147
Db 975 GCCTCGGCGCTCAGCGGACCCCTCTCGCCCTCGAGCACAAGCTCGGACACCCTGCTC 916
QY 148 AACCTAGACACGACGAGATTGCGCTGACCGAGCGGTAGATTACTACAGGGCTGC 207
Db 915 AACCGCACACCGGCGCATGAGTGTACCGGAGGAGCGCGTACTTCTCGACACAGCA 856
QY 208 CGTCTGATGCTTCATGAAGTGCAGGATGTTCTATGAGCAACTGTATGCTTCAATAACACC 267
Db 855 CGGCAGATCTGACGAGATGGGCGCTCGAGGAACGCTTGGCCCTCGCGCGGAGAGC 796
QY 268 CCCATCGGACGCTAGCGATTGCTGTTTCAACTATGCGCAAAATGTTCTCGCGGG 327
Db 795 CCGCGCGGCGCTGCGGATCAACGCGGCTCGCCCTTCAATGCTGACGCGGTGTGCGG 736

QY 328 CTGACAGCAAAATGCTGAAAGATACCCAGGTTTGAGCGTCAATCTGTTACCGGAAT 387
Db 735 CACATCGGCGAGTTCGCGCGGTGTATCCCGGACATCCAGCTCGAATCAACACCAACGAC 676
QY 388 CCAGCCCCCGACTGATTCGCGACGCTCTGAGTGTGGTGTGATCCGCGTCGCGCGGTTGACG 447
Db 675 CTGATCATGACCTGTTGAGCAGCAGCAGCGGTGCGCATCGCATCGTGTGCGCTCAGC 616
QY 448 GATTTCAGGCTGTTTTCGCGCGCTCTGCGGCGATGCCAATGGTGTGCGCGCGGAAA 507
Db 615 GACTCCACCTGACGCGGCGCGCTGGGCGCGAGCGGCTGCACATCTCGCCACGCGCG 556
QY 508 AGCTATCTCACAAATACGCGCATACCGGAAACCGCGGATTTGAGTAGTCAATTCATGG 567
Db 555 GAGTACTGCGCGCGCACGCGCGCGGAGCATCGAGGGCTGCGAGGCGACACCCCTG 496
QY 568 CTGGAATACAGCGTGC 583
Db 495 CTCGGCTTACCAGC 480

RESULT 14

US-09-252-991A-13887
; Sequence 13887, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 13887

; LENGTH: 1809

TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13887

Query Match 11.4%; Score 106.4; DB 4; Length 1809;

Best Local Similarity 49.5%; Pred. No. 6.4e-24;

Matches 275; Conservative 0; Mismatches 281; Indels 0; Gaps 0;

QY 28 TTGCGCAAGTAGTGAATTTGCTCTTTTACCGCGCGCGCGGACAGTACAGATGAGC 87
Db 11 TTGCTCAGTGTAGTCACACCGGTTCCATCAGCGCGCGCGGAAACAGATGGAGCTGACC 70
QY 88 GTTTCGTCATCAGTCAGACGGTATCAAAATCGGAAGATGAGTTGACGTTAAAGCTGTTA 147
Db 71 GCCTCGGCGCTCAGCGGACCCCTCTCGCCCTCGAGCACAAGCTCGGACACCCTGCTC 130
QY 148 AACCTAGACACGACGAGCATTGCCTGACCGAGCGGTAGAAATTTACTACAGGGCTGC 207
Db 131 AACCGCACACCGGCGCATGAGCTGACCGAAGAGGCGGCTACTTCTCGACAGGCA 190
QY 208 CGTCTGATGCTTCATGAAGTGCAGGATGTTCTATGAGCAACTGTATGCTTCAATAACACC 267
Db 191 CGGCAGATCTGACGAGATGGAGGCGCTCGAGGAACGCTGCGCCCTGCGCGCGGAGAGC 250
QY 268 CCCATCGGACGCTAGCGATTGCTGTTTCAACTATGCGCAAAATGTTCTCGCGGG 327
Db 251 CCGCGCGGCGCTCGGATCAACGCGGCTCGCCCTTCAATGCTGACGCGGTGTGCGG 310
QY 328 CTGACAGCAAAATGCTGAAAGATACCCAGGTTTGAGCGTCAATCTGTTACCGGAAT 387
Db 311 CACATCGGCGAGTTCGCGCGGTGTACCGGACATCCAGCTCGAATCAACACCAACGAC 370
QY 388 CCAGCCCCCGACTGATTCGCGACGCTGTGGATGTGCTGATCGCGGTGTGCGG 447

Db 371 CTGATCATGACCTGTTGAGCAACGACCGAGCTGGCGATCCGATCGGTGCGTCCAGC 430
QY 448 GATTCCAGCTGTTTCCCGCGCTCTGGGGGAGTCCCAATGGTGGTGTGCGCCCGGAAA 507
Db 431 GATCCACCTTGACCGCGCCGCTGGGGGCGAGCGGCTGCACATCTCTGCCAGCCCG 490
QY 508 AGCTATCTCACAAATACGAGCATACCGGAAACCGCGGATTTGAGTAGTCAATTCATGG 567
Db 491 GAGTACCTGGCCGCGACCGCGCGGAGCATCGAGGGCTGGAGGGGACACCCCTG 550
QY 568 CTTGAATACAGGTGC 583
Db 551 CTCGGCTTACCAGC 566

RESULT 15

US-09-489-039A-454
; Sequence 454, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 454
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-454

Query Match 11.1%; Score 103.2; DB 4; Length 933;
Best Local Similarity 49.3%; Pred. No. 4.8e-23;
Matches 270; Conservative 0; Mismatches 278; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCAATGTCGGTGTTCGCAAGTAGTGAATTTGGCTCTTTTACC 60
Db 28 ATGGACCGGTTTTCAGCTCTCAAGCCCTCACCGGTTGTGAAAGCGCGAGCTTTACC 87
QY 61 GCGCGCGCCAGACAGATGAGCGTTTCGTCATCAGTCAGACCGGTATCAAAACTG 120
Db 88 CGCGCGCGGAGTCGCTGAATATGCCAACGCCACCCCTCAGCAAAACCATTCAGCAACTG 147
QY 121 GAAGATGAGTTCAGGTAAAGCTGTTAAACCGTAGCACACGAGCAATGGCTGTACCGAA 180
Db 148 GAAGGCACTTTGGGTTTCCCTCCTGAGCGCACCGCGCGTATTACCGTGACGCGG 207
QY 181 GCGGTAGAAATTACTACAGGGCTGCGCTGTATGCTTCATGAAGTGCAGGATGTTCAAT 240
Db 208 GAAGGACGGGAGTACTATGAAAAGCCCGCTGTCTGTGGAAGACCTGGAAGAGATCGAC 267
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGAGCGTACGCAATGGCTGTGTTTCA 300
Db 268 GCGCTTTCAATACCGCCGCAATAGCGAGGCGCATCTACGATCGCATCGGCGGG 327
QY 301 ACTATGGCAAAATGTTCTCCGCGGTGACAGCAAAATGCTGAAAGATPACCCAGGT 360
Db 328 TCGACCGCGTGGGATGTGTGATCCGCTGTCTGGGAGCTTTATGACCTCTGGCGGGAT 387
QY 361 TTGAGCGTCAATCTGGTTACCGGAATCCAGCCCGGACCTGATTGCCGACGGTCTGGAT 420
Db 389 ATCCGTATCGATTGAGGTGCCGATAGCCCGGACCTCATCAGCGCAATATTGAC 447
QY 421 GTGGTATCCGCGTGGCGGCTTGAGGATTCAGGCTGTTTCCCGCGCTCTGGGCGG 480
Db 448 TCGGCATCCGCGGCTCGATGGAAGACTGACGCTGATCGCCCGTAAATCGCGAG 507
QY 481 ATGCCAATGGTGTGGCCCGGAAAGAGCTATCTCACAAATACGGCATACCGGAAAAA 540

Db 508 GCGACGCTGGTCACTTGGGCAACCCCGGGCTATCTTCAGGCTACGGCACCCCGCCTCG 567
QY 541 CCCGCCGA 548
Db 568 CCGGATGA 575

Search completed: December 18, 2004, 20:48:55
Job time : 108.036 secs

This Page Blank (uspto)

Db	61	GCCGCCGCCAGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGTATCAAAACCTG	120
Qy	121	GAAGATGAGTTGCAGGTAAAGCTGTAAACCGTAGCACACGACAGCAATTGGCTGACCGAA	180
Db	121	GAAGATGAGTTGCAGGTAAAGCTGTAAACCGTAGCACACGACAGCAATTGGCTGACCGAA	180
Qy	181	GCCGGTAGAATTTACTACACAGGGTGCCTGTGATGCTTCATGAAGTCAGAGATGTTCAAT	240
Db	181	GCCGGTAGAATTTACTACACAGGGTGCCTGTGATGCTTCATGAAGTCAGAGATGTTCAAT	240
Qy	241	GAGCAACTGTATCGCTTCCAATAACACCCCAATCGGACGCTACGCAATGGCTGTTCTTCA	300
Db	241	GAGCAACTGTATCGCTTCCAATAACACCCCAATCGGACGCTACGCAATGGCTGTTCTTCA	300
Qy	301	ACTATGGCACAAAATGTTCTCGCCGGCTGACAGCCAAAATGCTGAAGAATACCCAGGT	360
Db	301	ACTATGGCACAAAATGTTCTCGCCGGCTGACAGCCAAAATGCTGAAGAATACCCAGGT	360
Qy	361	TTGAGCGTCAATCTGGTTACCGGAATTCACAGCCCCGACCTGATTCGCAGCGTCTGGAT	420
Db	361	TTGAGCGTCAATCTGGTTACCGGAATTCACAGCCCCGACCTGATTCGCAGCGTCTGGAT	420
Qy	421	GTGGTGATCCGGCTCGCGCGTTGCAGGATTCACAGCCTGTTTCCCGCCGCTCTGGCGCGG	480
Db	421	GTGGTGATCCGGCTCGCGCGTTGCAGGATTCACAGCCTGTTTCCCGCCGCTCTGGCGCGG	480
Qy	481	ATGCCAATGGTGTGGCGCGCGAAAGCTATCTCAACAATACGGCATACCGGAATA	540
Db	481	ATGCCAATGGTGTGGCGCGCGAAAGCTATCTCAACAATACGGCATACCGGAATA	540
Qy	541	CCCGCGAATTCAGTAGTCAATCATGCGTCTGAATACAGCGTGGCGCCGCAATGAATTT	600
Db	541	CCCGCGAATTCAGTAGTCAATCATGCGTCTGAATACAGCGTGGCGCCGCAATGAATTT	600
Qy	601	GAATGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAACAAGGAATTTTGACT	660
Db	601	GAATGATCGCACCGGAAGGATCTCGACTCGCTGATCCCAACAAGGAATTTTGACT	660
Qy	661	AATGATCCGATGACACTGGTCGTGGCTGACGGCGGCTGCCGGGATCGCTACGTGCCG	720
Db	661	AATGATCCGATGACACTGGTCGTGGCTGACGGCGGCTGCCGGGATCGCTACGTGCCG	720
Qy	721	CTGATGTGGTGCATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCCGCGTTAC	780
Db	721	CTGATGTGGTGCATCAACGAGATCAATCGTGGGAGCTGGAGATCTGCTGCCGCGTTAC	780
Qy	781	CAGTCAGATCCACGCCCGGTTATTCGTTATATACCGAAAAAGATAAGCTCCGCTGAAG	840
Db	781	CAGTCAGATCCACGCCCGGTTATTCGTTATATACCGAAAAAGATAAGCTCCGCTGAAG	840
Qy	841	GTACAGGTGCTGATCAATCGCTGACGAGTATTTTGTGAGGTCGGTAAATTTGTTTCAG	900
Db	841	GTACAGGTGCTGATCAATCGCTGACGAGTATTTTGTGAGGTCGGTAAATTTGTTTCAG	900
Qy	901	GAGATCCACGGCGCGGGAAGAGAGTAA	930
Db	901	GAGATCCACGGCGCGGGAAGAGAGTAA	930

RESULT 2
US-10-8282-122A-11996
; Sequence 11996, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant

```

/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/257,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 11996
/ LENGTH: 939
/ TYPE: DNA
/ ORGANISM: Burkholderia cepacia
/ US-10-282-122A-11996

```

Query Match	12.18;	Score 112.4;	DB 15;	Length 939;
Best Local Similarity	47.8;	Pred. No. 1.5e-27;		
Matches 394;	Conservative	0;	Mismatches 421;	Indels 9; Gaps 2;
QY	1	ATGGAACGACTAAACGCATGTCGGGTGTTGGCCAAAGTAGTGAATTTGGCTCTTTTACC	60	
Db	1	ATGGATACGTTTACAAACATCGGGGTATTGTCGGCGTGGTCGAGCGCGGCGAGCTTTACC	60	
QY	61	GCGCGCGGCACAGCTACAGATGAGGGTTTGGTCCATCAGTCAGACGCTATCAAACTG	120	
Db	61	GCGCGCGGCCACGAGATGAATTCGACCGCCTACGGCTCGCGCGGTCTCGGATCTC	120	
QY	121	GAAAGATGAGTTGCGAGGTAAAGCTGTTAAACCGTAGCACCGCAGCATTTGGCTGACCGAA	180	
Db	121	GAGGCCACCTGCGCACCGGCTCTCTGAACCGCACGCGCGCGGATCCGGCTGACCGAG	180	
QY	181	GCGGTGAAATTTACTACGAGGCTCGCGTGGTATCTTTCATGAAGTCAGAGATGTTTCAT	240	
Db	181	GCGCGCGAGCGTTATCTGACGCGTGGGAAACAGATCCTCGCTTAGCTGCACCGAGCCGAA	240	
QY	241	GAGCAACTGTATGCTTCAATAACACCCCCCATCGGACGCTACGATTTGGCTGTTCTTCA	300	
Db	241	GCGGAGGCGGCGACGGCGACCGCGCGCGCTCGCGCAAGCTGAAGTCCATTGCTTCAG	300	
QY	301	ACTATGGCACAAATGTTCTCGCGGGCTACAGCCAAAATGCTGAAGAAATACCAAGT	360	
Db	301	AGCCTCGGCCAGCACTATCTGGTGGCCGCCATCGCGGCTACCGGAGCGCTATCCGAC	360	
QY	361	TTGAGCGTCAATCTGTTTACGGAAATTCAGCCCCGACCTGATTCGCCAGCGTCTGGAT	420	
Db	361	GTGACGTCGAGCTGACGCTCGCGCAGCGGATCCCGACCTGCTCGACGAGGGGTACGAC	420	
QY	421	GTGGTATCCGGCTCG--GGCGTTGCAGGATTCAGGCTGTTTCCCGCGCTCTGGGC	477	
Db	421	GTCCGGATGTCGTGCGCCGCGACCTGCCGATTCGGGCTCGTGTGCGAGCGGCTCGGC	480	
QY	478	GCGATGCCAATGGTGGTGTGCGCGCGGAAAGCTATCTCACAAATACGGGATACGGAA	537	

Db	481	GAGAGTACAGCGTGGTCTGCGCATCGCGGGCTACGTCGAGTCGACCGCGTGC	CGCGAG	540
Qy	538	AAACCCGCGGATTTGAGTAGTCATTATCGGTCTGAATACACGCTGGCGG-----	CCCGAC	591
Db	541	CGCGCGCGCATCTCGCACAGCACTGTGCTCGGGATGTCGCGCGGGCTTTTCA	TTTC	600
Qy	592	AATGAATTTGAACTGATCGCACCGGAAGGATCTCGACTCGCTGATCCACAAGGA	AGA	651
Db	601	GACCAATGGCGCTGGCGGGCGCAACCGCGACGAGTCTGCCGATCACGGCGCGCGG	CG	660
Qy	652	TTTGTACTAATGATCCGATGACGCTGGTGGCTGGCTGACGGCGGCTGCGGGATCGCG	TC	711
Db	661	TTTCGGGTGAACGTGCGCGAGCGCTCGCGTGGCCGTGCGGAGGGGATGGCGCTCGCG	TC	720
Qy	712	TACGTCCGCTCATGTGGGTGATCAACAGAGATCAATCGTGGGAGCTGAGATCCTGCTG	TC	771
Db	721	GGCGTCCGCTCTATTCCGCAGTCCGTGCGCTAGCGGGACATCTCGTGC	CGTGA	780
Qy	772	CCGCGTTACCATCGATGATCCACGCCGGTTTATCGTTATATAC		815
Db	781	CCGAGTACCGGTCCAGTGATGAACATCTACGCGTGTATCC		824

RESULT 3

```

US-10-282-122A-32141
; Sequence 32141, Application US/10382122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

Query Match	11.8%;	Score 109.6;	DB 16;	Length 945;
Best Local Similarity	50.7%;	Pred. No. 1.4e-26;		
Matches	290;	Conservative	0;	Mismatches 279; Indels 3; Gaps 1;

Qy	1	ATGGAACGACGTAAACCGCATGTCGGTCTTTGCCAAAGTAGTTGNAATTGGCTCTTTTACC	60
Db	1	ATGGAACCTTGCAAAACATCGTGTCTTCAGTTGCGTAGCCCAACTCGCGACGCTTCACT	60
Qy	61	GCGCCGCCAGACAGCTACAGATGAGCGTTTCGTTCATCAGTCAGACGGTATCAAAACTG	120
Db	61	GCCTCTGGCGGCAACTGGATAGACACCGGAACGTGTCGGGGCGGTCTCCAACTG	120
Qy	121	GAGATGAGTTGCGAGTAAAGCTTTAAACCGTAGACACGACGACATTGCGCTGACCGAA	180
Db	121	GAAGCCCATCTGCAAAACAGGCTGCTCAACCGTACCAACCGCGCATTTGCGCTGACCGAA	180
Qy	181	GCCTGTAGAATTTACTACACGAGGTCGCGTGTATGCTTTCATGAAGTCAGGATGTTTCAT	240
Db	181	GCCTGCAAGCGTTTACCTGATGCGTTGGGAACAGATTCTTACCTACGTCGAAGACCGGAG	240
Qy	241	GAGCACTGTATGCCCTCAATAACACCCCATCGGAGCGTACGCATTGGCTGTTCTTCA	300
Db	241	GCGGAAGCCAGACGCGCCATGCGCGCCCGCGCGGCGAGTTGAAGTGCATTGATGACT	300
Qy	301	ACTATGCAACAAATGTTCTTCGCGGGGCTGACAGGCCAAATATGCTGAAGAATATCCACAGT	360
Db	301	GGGTGCGGCAGCACTTCGTGGTTCGATGCCATCGCCGCTACCGGAACGACCCCGGAC	360
Qy	361	TTGAGCGTCAATCTGTTTACCGAATTCAGCCCCCGAAGCTGATTCGCGACGGTCTCGAT	420
Db	361	GTGACCTTCGATTGACCATGGCCAAACCGCGTCCCGATCTGCTCGACGAGGGTTATGAC	420
Qy	421	GTGGTGATC---CGCGTCGGCGGTTTCAGAGATTCAGAGCCTGTTTTCCCGCGCTCGGGC	477
Db	421	GTCTCCATGCTGCTGGCCACCGAATCGCCGACTCGGGGTTCGTTTCGACAGCGCTGGC	480
Qy	478	GCATGCCAATGTGTGTGTCGGCGGGAAGTATCTTCAACAATACGCGCATACCGGAA	537
Db	481	ATCACTTACAGCAATTGCTGTGCTCGCCGCTACATCGCCAGCATGCGTGGCGCAC	540
Qy	538	AAACCCGCGCATTTGAGTAGTCATTATGGCT	569
Db	541	AAGCTTGGCGACCTGCTCAAGACGCTGCT	572

RESULT 4	
US-10-282-122A-12996	
; Sequence 12996, Application US/10282122A	
; Publication No. US20040029129A1	
; GENERAL INFORMATION:	
; APPLICANT: Wang, Liangsuo	
; APPLICANT: Zamudio, Carlos	
; APPLICANT: Malone, Cheryl	
; APPLICANT: Haselbeck, Robert	
; APPLICANT: Chlsen, Kari	
; APPLICANT: Zyskind, Judith	
; APPLICANT: Wall, Daniel	
; APPLICANT: Trawick, John	
; APPLICANT: Carr, Grant	
; APPLICANT: Yamamoto, Robert	
; APPLICANT: Forsyth, R.	
; APPLICANT: Xu, H.	
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms	
; FILE REFERENCE: ELITRA.034A	
; CURRENT APPLICATION NUMBER: US/10/282,122A	
; CURRENT FILING DATE: 2003-02-20	
; PRIOR APPLICATION NUMBER: 60/191,078	
; PRIOR FILING DATE: 2000-03-21	
; PRIOR APPLICATION NUMBER: 60/206,848	
; PRIOR FILING DATE: 2000-05-23	
; PRIOR APPLICATION NUMBER: 60/207,727	
; PRIOR FILING DATE: 2000-05-26	

PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12996
LENGTH: 936
TYPE: DNA
ORGANISM: Burkholderia fungorum
US-10-282-122A-12996

Query Match 11.4%; Score 105.6; DB 16; Length 936;
Best Local Similarity 46.7%; Pred. No. 3.3e-25;
Matches 410; Conservative 0; Mismatches 459; Indels 9; Gaps 2;

QY 1 ATGGACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 1 ATGGATACGCTTCAAAACATGCGGTATTTGCTGTCTGTCGAGCGGACGTTTACG 60
QY 61 GCGCGCGCCAGACGATACAGATGACGCTTTCGTCCATCAGTCAGACGCTATCAAACTG 120
DB 61 GGTGCGCGCAGCATCTGAACACGACACCGGCTATCGTCGCGCGGTGTCGATCTG 120
QY 121 GAAGATGAGTTGCGAGTAAAGCTTTAAACCGTAGACACGACGATTTGGCTGACCGAA 180
DB 121 GAAGCGCATCTGCGACCGGACGCTGTAACCGACACCGCGCGGATTCGCTGACCGAG 180
QY 181 GCGGTGAGAAATTTACTACACGAGGCTCCGCTGCTATGCTTCATGAAGTGCAGGATTTCA 240
DB 181 GCGCGGCGGCTATCTGACGCGCTCGAGCAGATTTCTCGGTACGTGACCGCGGAA 240
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGATTTGGCTGTTCTTCA 300
DB 241 GCGGAGGCGGCGACGCCCATCGCTCGCTCGGCAAGCTCAAGGTCCACGCCATGACG 300
QY 301 ACTATGCGCAAAATGTTCTCGCGCGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360
DB 301 AGTTTCGCGCAGCACTATGTGTGTCGCGCTGTGTCGCGCTATCAGACGCTTACCGGAC 360
QY 361 TTGAGCGTCAATGTGTTACCGGAATTCAGGCCCGGACCTGATTCGCGACGCTGAT 420
DB 361 GTCCATATCGAACTGACGCTCGGCGACGCTGCGCGATCTGCTGACGAGGCTTCGAC 420
QY 421 GTGGTATC---CGCTCGCGGCTGACAGGATTCAGCGCTGTTTCCCGCGCTGCGG 477
DB 421 GTGTGCTACGCTCGCAACCGGTTTCCGGAATTCGGGCTGCTGCTGCGAGGCTTCGCG 480
QY 478 GCGATGCAATGTGTGTGTCGCGCGGAAAGCTATCTCACAATACGCGATACCGGAA 537
DB 481 AGCGGCTTACGATTCGATGCGGCTCGCGGCTATCTGAGCTGACGCGCTGCGGAA 540
QY 538 AAACCCGCGATTTGAGTAGTCAATTCATGCTGTGATACAGGTGCGG-----CCCGAC 591
DB 541 ACGCCGCGGACCTTCGCGGTACACTGTCTGCAAAATGTCACGCGGTTTTCGCGAC 600
QY 592 AATGAATTTGAATGATCGCACCGGAGGAGTCTCGACTCGGCTGATCCCAAGCAAGA 651
DB 601 GACAAATGACCTTCGACGCGCCGACGCGGAGGAAACGATCGCCCTCGGACCGACCT 660
QY 652 TTTGTGACTAATGATCGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 711

661 TTCCAGGTGAAGCTGCGCGAAGCGATGCGGTGCGCGTATCGAGCGGCATGCGCGTCGGG 720
QY 712 TAGCTGCCGCTGATGTTGGGTGATCAACGAGATCAATCGTGGGAGCTGAGATCTTCTGTG 771
DB 721 CTGATTCGATCTACTCGCGGATCAGCGCTTGCAGCGGAGAACTGGTGTGGCTGCTG 780
QY 772 CCGCGTTTACAGTCAGATCCACGCGGTTTATGCGTTATATACCGAAAAAGATAAGCTG 831
DB 781 CCGGAATACAGTCCGCGAGGATGAATCTGTACGCGTTGTACCCCTCGCGCGAGTATCTG 840
QY 832 CCGCTGAAGGTACAGTCTGTCGATCACTCGCTGACCGA 869
DB 841 GACGCGAAGATTGCGACGTGGATCGAGTTCTCTGCGCGA 878

RESULT 5
US-10-127-032-41
; Sequence 41, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-41

Query Match 10.7%; Score 99.8; DB 15; Length 918;
Best Local Similarity 49.4%; Pred. No. 3.3e-23;
Matches 288; Conservative 0; Mismatches 292; Indels 3; Gaps 1;

QY 2 TGNACGACTAAACGCGATGCTGCTGTTTCCAAAGTAGTTGAATTTGGCTCTTTTACCG 61
DB 14 TGGATCACCTCACCGCTCTCAAGGTCTTCGCGCGCTGCGCGGAAATGCGGCTTCGCG 73
QY 62 CCGCGCGCAGACGCTACAGATGAGCGTTTTCGTCCATCATGTCAGACGCTATCAAACTGG 121
DB 74 CCGCGCGCGCGAGATGATCTCTCGCGCGCGCGCTGAGCAAGAACGTCGCGAGCTG 133
QY 122 AAGATGAGTTGCGAGTTAAGCTGTTAAACCGTAGCACACGACGATTCGCTGACCGAAG 181
DB 134 AAGCGCACCTCAAGGTGCGCTGATCAATCGCACCCACCGCGAGCTGAGCTTACCGAGG 193
QY 182 CCGGTAGAAATTTACTACAGGCGTTCGCTGATGTTTTCATGAAGTCGAGATGTTTCATG 241
DB 194 CCGCGGAGTCTACCGCGACGCGCTGGAGGCGATCTCTGACGACCTCGAGCGCGCGACG 253
QY 242 AGCAACTGTATGCTTCAATAACACCCCATCGGAGCTACGATTCGCTGTTCTTCAA 301
DB 254 CCGCGCTCACTTCGATGACGAGCGGCGCGCGCTGCTGCGGCTGACGCGCGCGCTGA 313
QY 302 CTATGCGCAAAATGTTCTCGCGGCTGACAGCAAAATGCTGAAAGATACCCAGTT 361
DB 314 CCCTCGCGCTCACCTGCTGACCCCGGCTATTCGGCTTTTCTCCAGCGTTATCCGAGC 373
QY 362 TGACGCTCAATCTGTTTACCGGAATTCAGCCCGCGACCTGATTCGCGAGCTTCTGATG 421
DB 374 TGGCGCTGAACTGCTCTCTCGAGGACGCGCGCGCGAGCACTGATCGCGAAGGATCGACC 433

QY 422 TGGTATCCGCG---TCGGCGGCTTGCAGGATTCAGCCTGTTTCCCGCCGCTCTGGCG 478
Db 434 TGGCCCTGCTGTGTAGCGACCGGTTTGCAGCTCCGGCTGTTCGCCCGCCGCTGCTGG 493
QY 479 CGATGCCAATGTGTGTGCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAA 538
Db 494 TCTTGAGACAGCTGCTCTCGCGGCGCGGCTTAOCTCAGTCAGATGCGCCAGCGCTGC 553
QY 539 AACCCCGCGATTGTAGTAGTCTATTCATGGCTTGAATACAGCGT 581
Db 554 GGCCGGAGGCCCTGCGCGACACAGAGTGATCCGCTTACGCT 596

RESULT 6

US-10-282-122A-13932
; Sequence 13932, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 13932

; LENGTH: 939

; TYPE: DNA

; ORGANISM: Burkholderia mallei

US-10-282-122A-13932

Query Match 10.4%; Score 96.4; DB 16; Length 939;
Best Local Similarity 45.5%; Pred No. 5, 1e-22;
Matches 426; Conservative 0; Mismatches 50; Indels 9; Gaps 2;

QY 1 ATGGAACAGCTAAACCGATGCGGTGTGTTGCCAAAGTAGTGAATTTGGCTCTTTTACC 60

Db 1 ATGGAACAGCTCTCAAAACATCGCGTGTGTTGCCCGGTGGTCCGATGCGGGCAGTTTACG 60

QY 61 GCGCGCGCCAGACACTACAGTACGCTTCGTCCTACAGTACAGAGTATCAAACTG 120

Db 61 GCGCGCGCCAGACACTCAATTCCAGACCGCGGTACGCATCGCGCGGCTGTCGATCTC 120
QY 121 GAAGATGAGTTGAGGTAAAGCTGTTHAAACGTAGACACAGCAGCATTTGGCCTGACCGAA 180
Db 121 GAGCGCAGCTTTGGCGACCGCGTCTTCTGACCGGACGACGCGGCGCATTCGCGCTCACCGAG 180
QY 181 GCGCGTAGAATTTACTACAGGCTGCGCTGTATGTTTTCATGAAGTCAGGATGTTTCAT 240
Db 181 GCGCGGAGCGTATCTGCGAGCGCTGCGAGCAGATCCTCGGTACGTCGAGCAGCGGAG 240
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTTGGCTGTTTCA 300
Db 241 GCGGAGCAAGCGACGCGCACGCGCGCCCTCCGCAAGCTGAAAGTGCATCTCATGACG 300
QY 301 ACTATGGCACAAAATGTTCTCGCGGCTGACAGCCAAAATCTGTAAGAAATACCCAGGT 360
Db 301 AGCTTCGGCAGCACTACACGCTGCGCGCGGTGCGGCTACCGGACGCGCTATCCGAA 360
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCCCGACCTGATTTGCCGAGCTCTGGAT 420
Db 361 GTCCAGGTCGACCTGACGCTCGCGCAGCGGCTCGCGATCTGCTGACGAGGCTTCGAC 420
QY 421 GTGCTGATCCGCGTCCGCG---GCGTTGACGATTCAGCCTGTTTCCCGCGCTCTGGGC 477
Db 421 GTGCTGCTGCTGCTCGCGCGGAGCTGCGCGATTCGCGGCTCTGTCGACGCGCTCGGC 480
QY 478 GCGATGCCAATGTGTGCTGCGCGGCGGAAAAGCTATCTCACAAATACGCGCATACCGAA 537
Db 481 GAGAGCTTCAGCGTCTGCTGCGCGCTCGCGCGGTACATCGAGCAGCAGCGCGCGCGAG 540
QY 538 AAACCCCGCATTTGAGTAGTAGTATTCATGGCTTGAATACAGCGTG-----CGGCCGAC 591
Db 541 CGGCCGCGAGATCTTCCCGGCCACGCTCTGCTCGGGATGCTGCGCGCGGCTGTATGG 600
QY 592 AATGAATTTCAACTGATCGCACCGGAGGATCTCGACTCGCGCTGATCCACAGGAAGA 651
Db 501 GACGAATGGAAGTGTGTGGGCGCGCAGGCGCAGCAGCGTGACGCTCGCGCGCGCGCG 660
QY 652 TTTGTGACTAATGATCGATGACGCTGTGCTGCTGCTGACGCGGCTGCGCGGATCGCC 711
Db 661 TTTGCGGTGAACGTCGCGGAGCGGCTCGCGCGCGGATTCGCGAGGCGCATGGCATCGGC 720
QY 712 TACGTGCGCTGATGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTG 771
Db 721 GTGCTGCGGCTCTATTCGGGATCGCGCGGTTGCGCAACGCGGATTCGTCGCGGTGATG 780
QY 772 CCGGCTTACAGTCAGATCCACCGCGGTTTATGCTTATATACCGAAAAGATAAGCTG 831
Db 781 CCGGAATACCGGTGCGCACGCTGATGAACATCTACGCGCTCTACGCTCGCGCAGTATCTC 840
QY 832 CCGCTGAAGGTACAGGTGCTGATCACTCGCTGACGCGATATTTTGTGAGGTGCGTAA 891
Db 841 GACGCGAAGATCCGACGCTGGGTGCTATTTCTGCGGACGAGTTGCGCTGATCTCTCGAA 900
QY 892 TTGTTTCAGGAGATGACGCGCGCGGAAAGAGAAG 927
Db 901 GCGCAGCGCGCGGCTCGAGCGCTTCACCGGTGAG 936

RESULT 7

US-10-282-122A-33170

; Sequence 33170, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert		481		ATTACCTACAGCATGTTGTGTCATCCCCCGAGTAGTCGTCACAAACCTTCGGCATGGCGCAC	540
APPLICANT: Forsyth, R.		QY		538	AAACCCGCGATTGAGTAGTCATTATCGGT 569
APPLICANT: Xu, H.		Db		541	AAGCCGCGAGACCTTTTGAACCATGGTGCCT 572
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms		RESULT 8			
FILE REFERENCE: ELITRA 034A		US-10-282-122A-30384			
CURRENT APPLICATION NUMBER: 60/191,078		Sequence 30384, Application US/10282122A			
PRIOR FILING DATE: 2000-03-21		Publication No. US20040029129A1			
PRIOR FILING DATE: 2000-05-23		GENERAL INFORMATION:			
PRIOR FILING DATE: 2000-08-26		APPLICANT: Wang, Liangsu			
PRIOR FILING DATE: 2000-10-23		APPLICANT: Zamudio, Carlos			
PRIOR FILING DATE: 2000-11-27		APPLICANT: Malone, Cheryl			
PRIOR FILING DATE: 2000-09-06		APPLICANT: Haselbeck, Robert			
PRIOR FILING DATE: 2000-12-22		APPLICANT: Ohlsen, Karl			
PRIOR FILING DATE: 2001-02-09		APPLICANT: Zyskind, Judith			
PRIOR FILING DATE: 2001-02-09		APPLICANT: Wall, Daniel			
PRIOR FILING DATE: 2001-02-16		APPLICANT: Trawick, John			
Remaining Prior Application data removed - See File Wrapper or PALM.		APPLICANT: Carr, Grant			
NUMBER OF SEQ ID NOS: 78614		APPLICANT: Yamamoto, Robert			
SOFTWARE: Patent in version 3.1		APPLICANT: Forsyth, R.			
SEQ ID NO 33170		APPLICANT: Xu, H.			
LENGTH: 945		TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
TYPE: DNA		FILE REFERENCE: ELITRA 034A			
ORGANISM: Pseudomonas syringae		CURRENT APPLICATION NUMBER: US/10/282,122A			
US-10-282-122A-33170		CURRENT FILING DATE: 2003-02-20			
Query Match 10.1%; Score 93.6; DB 16; Length 945;		PRIOR FILING DATE: 2000-03-21			
Best Local Similarity 49.0%; Pred. No. 4.7e-21;		PRIOR APPLICATION NUMBER: 60/191,078			
Matches 280; Conservative 0; Mismatches 289; Indels 3; Gaps 1;		PRIOR FILING DATE: 2000-05-23			
QY		PRIOR FILING DATE: 2000-08-26			
Db		PRIOR FILING DATE: 2000-10-23			
QY		PRIOR FILING DATE: 2000-11-27			
Db		PRIOR FILING DATE: 2000-09-06			
QY		PRIOR FILING DATE: 2000-12-22			
Db		PRIOR FILING DATE: 2001-02-09			
QY		PRIOR FILING DATE: 2001-02-09			
Db		Remaining Prior Application data removed - See File Wrapper or PALM.			
QY		NUMBER OF SEQ ID NOS: 78614			
Db		SOFTWARE: Patent in version 3.1			
QY		SEQ ID NO 30384			
Db		LENGTH: 954			
QY		TYPE: DNA			
Db		ORGANISM: Pseudomonas aeruginosa			
QY		US-10-282-122A-30384			
Db		Query Match 10.1%; Score 93.6; DB 16; Length 954;			
QY		Best Local Similarity 49.0%; Pred. No. 4.8e-21;			
Db		Matches 280; Conservative 0; Mismatches 289; Indels 3; Gaps 1;			
QY		QY		1	
Db		Db		1	
QY		QY		61	
Db		Db		61	
QY		QY		61	
Db		Db		61	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	
Db		Db		121	
QY		QY		121	

QY 181 GCGGTAGATTACTACAGGCTGCGGTATGCTTTCATGAGTGCAGATGTTTCAT 240
Db |||||
QY 181 GCGGCGCAGCGTACTGCTGCGCTGCGAGAGATCTCCGCTAGCTCGAGAAAGCCGAG 240
Db |||||
QY 241 GAGCAACTGTATGCTTCAATAACACCCCGCTATCGGAGCGTACGCAATGGTGTTCATCA 300
Db |||||
QY 301 ACTATGGGACAAATGTTCTGCGCGGTGAGAGCAAAATGCTGAAGATATACCCAGGT 360
Db |||||
QY 301 GGCATCGGCGACGATACGTGATCGCGCCATCGCGCGCTACCGCCAGCATACCGGAC 360
Db |||||
QY 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCCCGACCTGATTCGCGAGCGGTCTGGAT 420
Db |||||
QY 421 GTGTGTATCCGCGT---CGGCGGTGAGAGATTCAGCCTGTTTTCGCGCTGCGG 477
Db |||||
QY 478 GCGATGCCAATGGTGTGCGCGCGGAAAGCTATCTCACACATACGGCATACCGGAA 537
Db |||||
QY 538 AAACCGCGGATTTGAGTAGTCAATTCATGGCT 569
Db |||||
QY 541 GAGCCTTCGAGCTGGCCAGCCAGCTGCCT 572

RESULT 9
US-10-282-122A-19983
; Sequence 19983, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19983
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-282-122A-19983

Query Match 9.7%; Score 89.8; DB 16; Length 896;
Best Local Similarity 47.2%; Pred. No. 9.5e-20;
Matches 342; Conservative 0; Mismatches 377; Indels 6; Gaps 2;

QY 49 GGCTTTTACCCCGCGCCAGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAAGC 108
Db 40 GGCAGCTTTTCGCGCGCGCGAGGAGCTGGCATGTCCGCCACCGCGGTTCAGCAACGC 99

QY 109 GTATCAAACTGAAGATGAGTTGCAGGTAAAGCTGTTAAACCGTAGCACACGCACATT 168
Db 100 ATCGCGCGGTGAGAGCGCCCTTGATACCGGCTTTTAAACCGACACCCCGCGGTG 159

QY 169 GGCCTGACCGAAGCCGCTAGAAATTTACTACAGGGCTGCGTGTCTGTCTTCAATGAGTG 228
Db 160 GCGCTACCGACCGCGGACAGCGCTACGTGGCGGTATCGGCCGCGCTGCAAGGATC 219

QY 229 CAGATGTTTCATGAGCAACTGTATGCTTCAATAACACCCCATCGGAGCGTACGCAAT 288
Db 220 CGTCTCGCCAGCGAGGATCCACAGCGACACCGGGGAAACCCCGCGCGCTGCGTCTG 279

QY 289 GGCTGTTCTTCAACTATGGCACAATAATGTTCTCGCGGGCTGACAGCCAAAATGCTGAAA 348
Db 280 AAGTGCAGAACCATATCGCGCGCTGTTCTGACAGCTGCTGATGACATTATGATC 339

QY 349 GAATACCCAGTTTGAGCGTCAATCTGTTACCGGAATTCAGCCCGCGAGCTGATGCC 408
Db 340 CGCTACCCGAAAGATGCGCGTAGAGACCGCTCAGCGAAGCGGAAATGATCGACATCGTCGG 399

QY 409 GACGCTCTGATGTGTGATCGCGCTGCGCGTTCAGGATTCAGGCTGTTTTCGCCG 468
Db 400 GAAGGTACAGCGCGGGATCCGCTTGAGAGTCCGCTCAGGACATGATCGCGGTG 459

QY 469 CGTCTG---GGCGGATGCGAATGGTGTGTGCGCGCGGAAAGCTATCTCACAAATAC 525
Db 460 CCGCTGACCGGGGAGATCCGCCAGTTGGTCAACGCCACCGCGGACTATTTTGCACGCCAC 519

QY 526 GGCATACCGGAAAAACCGCGGATTTGAGTAGTCAATCATGGCTTGAATACAGCGTGGC 585
Db 520 GGCATTCGCGAGCGCGGAGATCTGCTTTGCGATCAGGGGATGGCATGCGCATGGCC 579

QY 586 CCGGAC---AATGAATTTGAATCTGATCGCACCGGAAGGATCTCGACTCGCCTGATCCCA 642
Db 580 CACGGCGGATCTACCTTTGGGAGCTGCGCGCGGGGAAACGTAACGCCCTTTCGGTG 639

QY 643 CAAGGAAGATTGTGATTAATGATCCGATGACGCTGCTGCTGCTGCTGAGCGCGGTGCC 702
Db 640 CCGCGCGCTTTCGAGCTCCGATCTTTTTCGCTCGATCCGGCGGTAAGAGCGGATTA 699

QY 703 GGGATGCGCTACGTCGCGTGTGGGTGATCAACGAGATCAATCGTGGGAGCTGGAG 762
Db 700 GGGGTGGATTTTTCGCGGACTGTATATCCGGGAAGAGCTGGAGCTGGAGAGCTGGTG 759

QY 763 ATCCT 767
Db 760 AGCGT 764

RESULT 10
US-10-282-122A-13759
; Sequence 13759, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-03-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 13759

LENGTH: 936

TYPE: DNA

ORGANISM: Burkholderia fungorum

US-10-282-122A-13759

Query Match 8.4%; Score 78.4; DB 16; Length 936;
 Best Local Similarity 47.2%; Pred. No. 8.6e-16;
 Matches 272; Conservative 0; Mismatches 301; Indels 3; Gaps 1;

QY 1 ATGGACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
 DB 1 ATGGACGACTGGAGAACATCGGATGTTTCGTTCGAGTCGTGGAAGCAGCGAGCTTCACG 60
 QY 61 GCGCGCGGACAGCTACAGTAGCGGTTTGTCCTCATCAGTCAGACGGTATCAAACTG 120
 DB 61 AAGGCGCATCTCTTTTGTTCGGTGGCAGACCTCAGGTATCGCGTGCCATAACAGAACTG 120
 QY 121 GAAGATGAGTTCGAGTAAAGTGTAAACGCTAGCACACGAGCATTTGGCTGACCGAA 180
 DB 121 GAGTCGAGTTCGAGTAAAGTGTAAACGCGAGCGAGCAGACATGTCTACTAACCGAA 180
 QY 181 GCGGTAGATTTACTACAGGGTTCGGTGTATGCTTTCATGAGTGCAGGATGTTTAT 240
 DB 181 GCGGCGGACAGTTACTGCAAGTTGCGAACAGATTTCTTGAATACGTGCGTCTGCGCGAA 240
 QY 241 GAGCACTGTATGCTTCAATAACACCCCTCGGAGCGCTACGATTTGGCTGTTCTTCA 300
 DB 241 GCGGAGCGGAGGAGCGGCGGAGCGGCTCGCGGAAGCTGCGGATTCAGGACACACC 300
 QY 301 ACTATGGCAAAATGTTCTCGCGGGGTGACAGCCAAATGCTGAAGATATCCAGGT 360
 DB 301 AGTTTCGGGCAACATTATCTGCGGCCACTCATCGCGGCTATAACACGAGGATTCGCGAC 360
 QY 361 TTGAGCGTCAATCTGGTTACCGGATTCAGCCCGGACCTGATTGCGGAGGTCTGGAT 420
 DB 361 GTCTCTATGATCTTGTACTCGCAACAGGAGTGGCGGACATGATCGAGGAGGTTTGTAT 420

QY 421 GT---GGTATCGCGTCGGCGGTTGCAGGATTCAGGCTGTTTCCGCGGTCTGGGC 477
 DB 421 GTATCGGTGTCGTGCTAGAGAAATTGGCCGATTGGCACTCGTCTCACAACGCTATCGGT 480
 QY 478 GCGATGCCAATGGTGTGCGCGGAAAAAGTATCTCACAACAATACGCGCATACCGGAA 537
 DB 481 ACCACGTGCACGTGTTCTCTGTGCTTCGCCCCAGTACCTGTCGAGGAGGTATGCTGTC 540
 QY 538 AAACCCGCGGATTTGAGTAGTAGTATTCATGCTTGAA 573
 DB 541 TCACTAGAAGATCTGAACACGATACGTCGCTGCAA 576

RESULT 11

US-10-329-960-1/c

Sequence 1, Application US/10329960

Publication No. US20030099277A1

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.

TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Pragn

FILE REFERENCE: PB186P1

CURRENT APPLICATION NUMBER: US/10/329,960

CURRENT FILING DATE: 2003-01-02

PRIOR APPLICATION NUMBER: US 09/643,990

PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: US 08/487,429

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: US 08/426,787

PRIOR FILING DATE: 1995-04-21

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1830121

TYPE: DNA

ORGANISM: Haemophilus influenzae

FEATURE:

NAME/KEY: misc feature

LOCATION: (4747)..(4747)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (9921)..(9921)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (10150)..(10150)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (29298)..(29298)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (36543)..(36543)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (36551)..(36551)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (36636)..(36636)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (40808)..(40810)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc feature

LOCATION: (44416)..(44416)

OTHER INFORMATION: n equals a, t, g or c

FEATURE:

NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match 7.8%; Score 72.6; DB 14; Length 1830121;
Best Local Similarity 45.7%; Pred. No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
Query Match 7.88; Score 72.6; DB 16; Length 1830121;
Best Local Similarity 45.78; Pred. No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;
Qy 1 ATGGACGACTAAACGCGATGTCGGTGTTCGCAAGTAGTTCGAAATTTGGCTCTTTTACC 60
Db 1450229 ATGGATAAAGTCAACGCGATTTCTATTTCTGCAAGTCTTCAACCCCAAGCTTTTACT 1450170
Qy 61 GCGCGCCGACAGCTACAGTACGCTTTCCTCCATCATCAGTCAGCGGTATCAAACTG 120
Db 1450169 TTGGCAGCAAGCAACAAATATTTCTGTGCGATGCAAGCAATATAGTTTCAATTA 1450110
Qy 121 GAAGATGAGTTGCGAGTAAAGCTGTAAACCGTAGCACACGCGAGCATTCCTCCACCGAA 180
Db 1450109 GAAGACATTTAAACACACGATTTTACACGTCACACAGAAAAATATGCTACCGAA 1450050
Qy 181 GCGGTAGATTTTACTACAGGCTGCGTCTGATGCTTTCATGAGTGCAGGATGTTTAT 240
Db 1450049 GCTGGAATGATGATTTATATCAACGCTGTCAAGGCAATTTTGTCTTATGATGAAGCAGAT 1449990
Qy 241 GAGCACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATTCGCTCTTCTCA 300
Db 1449989 TCAGCATTTACTCAGCTTACCAGTTTATACAGGAAATTTACTGATTCCTCGT 1449930
Qy 301 ACTATGGCACAATAATGTTCTCGCGGCTGACAGCCAAATGCTGAAAGATACCCAGGT 360
Db 1449929 GATTTGGTTTACTTTTATTTGACCAAAATTTACCCACTTTTATGCTTAAGCATCTCAT 1449870
Qy 361 TTGAGCGTCAATCTGTTACCGGAATTCAGCCCGGCTGATTCGCGCGTCTGGAT 420
Db 1449869 TTGCATATGAGTTGAATTTAATGATAGAAATGATTTATATCCGAGGCTATGAT 1449810
Qy 421 GTGATGATCGCGTTCGCGGCTTCAGGATTCAGGATTCAGGCTGTTTCCGCGCTGCGCG 480
Db 1449809 CTTGCTCTTCGTATGTTTATGGAAGATAGCTGCTGTTGTTTACGTAATAATAGGAAT 1449750
Qy 481 ATGCCAATGTTGTCGCGGCAAGAGTATCTCACAAATACGCGCATACCGGAAAAA 540
Db 1449749 ACTACTGTTCAATTTTGAGCGTCTGCTTATTTCTTGAACGAATGTTATTCACAAACG 1449690
Qy 541 CCGCGCGATTT 551
Db 1449689 CTTGATGATTT 1449679

RESULT 13
US-10-158-865-1/c

```
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; FILE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```



```
/ NAME/KEY: misc feature
/ LOCATION: (122167)..(122167)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (122336)..(122336)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131340)..(131340)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (131360)..(131360)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (139910)..(139910)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (140398)..(140398)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (142750)..(142750)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145058)..(145058)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145171)..(145171)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145942)..(145942)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147197)..(147197)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (150841)..(150841)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152500)..(152500)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152530)..(152530)
/
Query Match 7.8%; Score 72.6; DB 18; Length 1830121;
Best Local Similarity 45.7%; Pred No. 4e-12;
Matches 252; Conservative 0; Mismatches 299; Indels 0; Gaps 0;
QY 1 ATGGAACGACTAAACGATGCGGTGTTGCCAAGTAGTGTGAATTGGCTCTTTACC 60
DB 1450229 ATGGATAAAGTCAACGCGATTTCTATTCTGCAAGTCAITGAAACCAAGTTTACT 1450170
QY 61 GCCGCGCCGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGCGGTATCAAACTG 120
DB 1450169 TTGGCAGCAAGCAACAAATATTCTGTGGCGATGGCAAGCAATAGTTTCAATTA 1450110
QY 121 GAAGATGAGTTGACGTTAAAGCTGTTAAACCGTAGCACACGCAATGGCTGACCGAA 180
DB 1450109 GAAGAACATTAAACACGATTTGTACACGTCACACGAAAAATATGCTTACCGAA 1450050
QY 181 GCCGTAGATTACTACAGGCTGCGGTGATGCTTCTATGAGTCAGCATGTTTCAT 240
```

Db 1450049 GCTGGAATGATGATTATTATCAACGCTGTCAAGCATTTTGTCTGATTAGTAGAAGCAGAT 1449990
QY 241 GAGCAACTGTATGCCITTCATAATAACACCCCATCGGACGCTAGCAGTATGGCTGTCTTCA 300
Db 1449989 TCCAGCATTAAGTACAGTTACAGTTTCAATCAAGGAAATTTACTGATTTCCGTTCCCTCGT 1449930
QY 301 ACTATGACACAAAATGTTCTCGCGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGT 360
Db 1449929 GATTTTGGTTTACTTTTATGACCAAAATTTACCCACTTTTATGGTAAAGCATCTCAT 1449870
QY 361 TTGAGCGTCAATCTGTTTACCGGAATTTCCAGCCCGGACCTGATTCGCGACGCTCTGGAT 420
Db 1449869 TTGCATATTGAAGTTGAATTAATGATAAGAAAATCGATTTATTTCCGAAGGCTATGAT 1449810
QY 421 GTGGTATCGCGTCGCGGCTGCGAGGATTCAGGCTGTTTCCCGCCGCTCTGGGCGG 480
Db 1449809 CTTGCTCTTCTGTTTATGGAAGATAGTCTGTTGTTTTCACGTAATAATAGGAAT 1449750
QY 481 ATGCCAATGTTGTTGCGCGGCAAAAGCTATCTCACAAATACGCGCATACCGGAAAAA 540
Db 1449749 ACTACTGTTTCATTTTGCAGCGTGCCTTAATTAATCTTGAAACGATGTTTCCACAAACG 1449690
QY 541 CCGCGCAATTT 551
Db 1449689 CCGTGAATTT 1449679

RESULT 14
US-10-282-122A-13890
; Sequence 13890, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13890
; LENGTH: 912

```

; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-13890

Query Match      7.7%; Score 71.8; DB 16; Length 912;
Best Local Similarity 45.1%; Pred. No. 1.6e-13;
Matches 393; Conservative 0; Mismatches 467; Indels 11; Gaps 3;

QY 4 GAACGACTAAACGATGTCGGTGTTCGCCAAGTAGTTGAATTGGCTCTTTTACGCC 63
DB 10 GATCGCTCGCGGCTGCACGATATTCATGACGCTGTCACGACGCGCAATTTACCGCGT 69

QY 64 GCGCCAGACACTACAGATGAGCGTTTCGTCATCAGTCAGCGTATCAAACTGAA 123
DB 70 GCGCGCGCGGCTCGGCTTGACGCGCGCGGTGACCTTGGCATCGGCACTCGAA 129

QY 124 GATGAGTTGAGGTAAAGCTGTTAAACGCTGACACACGAGCAATGGCTGACCGAAGCC 183
DB 130 AGCGAGCTGAAAGTGAAGCTCTTCAGCGCAGCAGCGCGCGTGAAGCTGACGGAAGCC 189

QY 184 GGTAGATTTACTACAGGCGTCCGTCGTATGCTTCAAGTCAGAGTTCATGAG 243
DB 190 GCGCAGCGTATTACGACGCAACGCGCAGCGCAGCGCTACCGGCAAGTTCATCACT 249

QY 244 CAACTGTATGCTTCAATAACACCCCATCGGAGCGCTACGATTTGGCTGTTCTTCACT 303
DB 250 GCGCTGAACGACGCGCGCGACGAAACCGGAGGCGCTTTCGGATCAGCGCGTCACTC 309

QY 304 ATGGCACAATAATGTTTCGCGCGGTGACAGCGCAAAATGCTGAAGAATACCGAGTTTG 363
DB 310 GCGAAATCGCTGTCGACGCGCGACGCTGTCGCGAGTTCCTCCGCGCTACCGAAGTG 369

QY 364 AGCTCAATCTGTTTACCGGATTCACGCGCCCGACCTGATTCGCGACGCTGCTGATG 423
DB 370 AGCTCGAGATCGCTACGAGATCATCTGTCGATCGTCAAGGACGCTCGACGCC 429

QY 424 GTGATCCGCTCGCGCGTGTGAGATTCAGCGCTGTTTCCGCGCTGTCGGGCGGATG 483
DB 430 GGAATCCGCTCGCGACCGCGTGCAGCGCGGATGTTGGGCTGAAGATCGCGCGCG 489

QY 484 CCAATGGTGTGCGCGCGAA---AGCTATCTCACATATACGATACCGGAA 540
DB 490 CTCAGTGCCTGCTGTCGCGCGCGCGCTATCTCGCGACGCGGCTGCGCGCGAAG 549

QY 541 CCGCGGATTTGAGTAGTATTCATGCTTGAATACAGCGTGC-----GCGCGCAAA 593
DB 550 CTCGCGATCTCGCGCGCACATGTCATCCGCTTTCGTTCCGAGAGCGCGCGCTG 609

QY 594 TGAATTTGAATGATCGCACCGAAGGATCTCGACTGCTGATCCCAAGGAGATTT 653
DB 610 CACAAATGCGCGCTGCGCGACGCGCGCGCGCGCTGAGCTGCGTGC-CGCGCGCTT 668

QY 654 TGTGACTAATGATCCGATACGCTGGTGGCTGGCTGACGCGCGGTGCGGATCGCCTA 713
DB 669 CGTCACGCGATACGCGCGCGTGTATCGACCGCGCGCGCGCGCATCGCTT 728

QY 714 CGTCCGCTGATGTGGGTGATCAACAGATCAATCGTGGGAGCTGAGATCTCTGCTCC 773
DB 729 CGTGTTCATCGCGAGCGCATCGAGACGACCTGCGCGCGCGCGCTGCGCGAGTGTCT 788

QY 774 GGTTCACAGTCAGATCCAGCGCGCTTTATCGCTTATATACGAAAGATTAAGTCC 833
DB 789 GCGGCGGCTGCGGAGCGCTGCGCGCGATGCGCTTCTACTACGGAACCGCAAGCAGT 848

QY 834 GTTGAAGGTACAGGTGCTGATCAACTCGCTG 864
DB 849 GCGCGCGAAGTGCAGCGCTTATCGCGGTG 879

```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITEA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11194
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-282-122A-11194

Query Match      7.6%; Score 71; DB 16; Length 690;
Best Local Similarity 46.1%; Pred. No. 2.7e-13;
Matches 276; Conservative 0; Mismatches 320; Indels 3; Gaps 1;

QY 1 ATGGAACGACTAAACGCAATGTCGGTGTTCGCCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 1 ATGGAATCTGATTTCAAAAGCATGCGAGCGTTTTCGTCACGCTCGCGACGCGGGGACCTTCACC 60

QY 61 GCGCGCGCAGACAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGCGTATCAAACTG 120
DB 61 GAGACCGGCAACCGCTCGGCATCACGCGCGCGCTGAGCGCGTTCGCTCGCTC 120

QY 121 GAAGATGAGTTGAGGTAAAGCTGTTAAACCGTAGCAGCAGGAGTTCGCTGACCGAA 180
DB 121 GAGCATCGTCTCGGCATCGGTTGCTGAACCGCGCGCTGTCATCGCTGTCGCTGACCGAA 180

QY 181 GCGGTAGATTTTACTACAGGCGTCCGTCGTCATCTTCATGAAGTCAGGATGTCAT 240
DB 181 GCGGCGAGCGCTATCTGCGCGCGCGCGGACATCACGCTCGCTCGATCGGAGGAG 240

QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGAGCGCTACGCAATTCGCTGTTCTTCA 300
DB 241 CGGAGCGCGCGACGCGCGCGTGTCTCGCACCGCGCGCTGCGCGCGATTCGACGCGCC 300

QY 301 ACTATGCGCAAAATGTTCTCGCGCGGCTGACAGCGCAAAATGCTGAAGAATACCCAGGT 360
DB 301 AGCATCGCAAAATGTTCTGCTGATTCGCTGATTCGCGCAATTCAGGCGGCTTACCGGAC 360

```

RESULT 15

US-10-282-122A-11194
; Sequence 11194, Application US/10282122A
; Publication No. US20040029129A1

```
QY 361 TTGAGGTCAATCTGTGTACCGAATTCAGCCCCCGAAGCTGATTCGCCGACGGTCTGGAT 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 GTGAGGTTCACCTGACCTGCTCGCGCGGCACCTGCCGAAACTGGTGGCGGACAGCTACGAC 420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GT---GGTGATCCGCGTCCGCGGTTGTCAGGATTCAGAGCTGTTTTCCCGCGCTCTGGGC 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 GTCGCGGTGATCGCGATGCCGAGCTGCCGAGTCCGAGAGGTCCGATCGACGTCCGC 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 GCGATGCCAATGTGTGTGTCGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAA 537
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 CGGATCCGACGCTGCTGTGCGCGTCCCGCGTATCTCGATCCGCAAGGATGCCCGAT 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 AATCCCGCGATTGTGAGTAGTATTCATGCTTGATCAGCGTGGCGCCCGACATGA 596
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 TCGCCGCGCGCGTGGTTACCAATCGCTGGTGCAACTGGTCCGCGCCCGGTATCACGA 599
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 18, 2004, 23:07:12
Job time : 660.306 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 22:46:01 ; Search time 4048.28 Seconds
(without alignments)
8371.204 Million cell updates/sec

Title: US-10-759-889-1
Perfect score: 930
Sequence: 1 atggaacgactaaacacat.....ggcggggaagaagaagtaa 930

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_gsa1: *
8: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	459	49.4	838	5	BQ144074
C 2	448	48.2	811	5	BQ143765
C 3	440	47.3	452	4	BQ451124
C 4	187.8	20.2	205	4	B1406127
C 5	167.6	18.0	182	4	B1406112
C 6	136.8	14.7	165	4	B1406111
C 7	131.8	14.2	755	8	AF094919
C 8	105.8	11.4	703	8	B2341676
C 9	101.4	10.9	565	9	CG708961
C 10	98.2	10.6	1027	9	CNS01FUZ
C 11	95	10.2	1178	8	B25533247
C 12	90.4	9.7	761	9	CG999406
C 13	89	9.6	954	8	B2558369
C 14	87.6	9.4	1279	8	B2567908
C 15	87.2	9.4	741	9	CG899239
C 16	84	9.0	1169	8	B2579305
C 17	82.8	8.9	550	8	BH389688
C 18	81.2	8.7	626	7	C0540363
C 19	80.4	8.6	777	7	CN822930
C 20	77	8.3	632	7	C0535994
C 21	76.6	8.0	1287	8	B2573075
C 22	74.8	8.0	716	7	CN822664
C 23	74.8	8.0	886	7	CK408871
C 24	74	8.0	531	7	CF887195

C 25 74 8.0 548 6 CB852476 UI-CF-FNO
C 26 74 8.0 548 7 CF891359 UI-CF-FNO
C 27 73.4 7.9 1241 8 B2567267 pac82-164
C 28 72.4 7.8 795 9 CL694387 PRI0164a-
C 29 68.6 7.4 1021 8 B2554085 pac81-60
C 30 67.8 7.3 789 7 CN824816
C 31 67.2 7.2 566 4 CK329341 H8160F09-
C 32 67.2 7.2 585 7 BG063353 H3006B12-
C 33 67.2 7.2 585 7 CK334119 H3006B12-
C 34 67.2 7.2 828 9 CL659241 PRI0133C
C 35 66.8 7.2 866 8 B2678395 PUBDR69TD
C 36 65.4 7.0 1241 8 B2559127 pac82-164
C 37 65.2 7.0 815 7 CN822455
C 38 65.2 7.0 820 7 CO202387
C 39 65.2 7.0 1227 8 B2577725 msh2-5544
C 40 65 7.0 964 8 B2558727 pac8401.4
C 41 64 6.9 1123 8 B2564676 pac82-164
C 42 61.6 6.6 730 5 BM964257 UI-M-EQO-
C 43 61.4 6.6 1208 8 B2552364 pac81-60
C 44 61.2 6.6 948 8 B2578978 msh2-6091
C 45 60.6 6.5 806 8 B2551243 pac81-60

ALIGNMENTS

RESULT 1
BQ144074/c

LOCUS BQ144074 838 bp mRNA linear EST 24-APR-2002
DEFINITION NF032D03DT1F1027 Drought Medicago truncatula cDNA clone NF032D03DT
5', mRNA sequence.
ACCESSION BQ144074
VERSION BQ144074.1 GI:20281133
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; -
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 838)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.,
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula drought library
Unpublished (2000)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 838 Std Error: 0.00
Plate: 032 row: D column: 03
Seq primer: TCACACAGGAACAGCATATCAC.
Location/Qualifiers
1..838
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF032D03DT"
/tissue_type="plantlets"
/dev_stage="pooled timepoints"
/clone_lib="Drought"
/note="Vector: Lambda Zap; Contains a mixture of entire
plantlets harvested in a series of days-post-watering
timepoints."

FEATURES

source

Query Match

Best Local Similarity

Matches 459; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

49.4%; Score 459; DB 5; Length 838;

Fred. No. 5.8e-133;

0; Mismatches

0; Indels

0; Gaps

0;

0;

0;

0;

0;

0;


```

source
1..452
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF09A07DT"
/tissue_type="Plantlets"
/dev_stage="Pooled timepoints"
/clone_lib="Drought"
note=vector: Lambda Zap; Contains a mixture of entire
timepoints. "

ORIGIN
Query Match 47.3%; Score 440; DB 4; Length 452;
Best Local Similarity 99.8%; Pred. No. 4.9e-127;
Matches 451; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 478 GCGATGCCAATGTGGTGTGCGCGCGCAAAAGCTATCTCACACAATACGGCATACCGGAA 537
DB 452 GCGATGCCAATGTGGTGTGCGCGCGCAAAAGCTATCTCACACAATACGGCATACCGGAA 393
QY 538 AACCCGCCGATTTGAGTAGTCATTCATGGCTTGAATACACGGTGCGCCGCAATGAA 597
DB 392 AACCCGCCGATTTGAGTAGTCATTCATGGCTTGAATACACGGTGCGCCGCAATGAA 333
QY 598 TTTCAACTGATCGCACCGGAGGAGTCTCGACTCGCTGTATCCACACGAGGAAGATTTCG 657
DB 332 TTTGAACTGATCGCACCGGAGGAGTCTCGACTCGCTGTATCCACACGAGGAAGATTTCG 273
QY 658 ACTAATGATCGATGACGCTGGTGCCTGTGCTGACGCGGGTGCGGGATCGCCTACGCTG 717
DB 272 ACTAATGATCGATGACGCTGGTGCCTGTGCTGACGCGGGTGCGGGATCGCCTACGCTG 213
QY 718 CCGCTGATGTGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTGCTCCGCGGT 777
DB 212 CCGCTGATGTGGTGATCAACGAGATCAATCGTGGGAGCTGGAGATCTCTGCTCCGCGGT 153
QY 778 TACAGTCAGATCCACGCCCGGTTTATGCGTTATATACCGAAAAAGATAAGCTCCCGCTG 837
DB 152 TACAGTCAGATCCACGCCCGGTTTATGCGTTATATACCGAAAAAGATAAGCTCCCGCTG 93
QY 838 AAGTCAGATC-GTGATCACTCGCTGACGGATTTTGTGAGTCGGTAAATTGTT 896
DB 92 AAGTCAGATC-GTGATCACTCGCTGACGGATTTTGTGAGTCGGTAAATTGTT 33
QY 897 TCAGGAGATGCACGGCGCGGGAAGAGAAGT 928
DB 32 TCAGGAGATGCACGGCGCGGGAAGAGAAGT 1

RESULT 4
BI406127/c
LOCUS
DEFINITION
155C0 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
linear EST 11-SEP-2002
sequence.
ACCESSION
BI406127
VERSION
BI406127.1 GI:15185541
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
tags
1 (bases 1 to 205)
REFERENCE
Crookshanks,M., Emmersen,J., Welinder,K.G. and Nielsen,K.L.
AUTHORS
The potato tuber transcriptome: analysis of 6077 expressed sequence
TITLE
tags
JOURNAL
FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE
21475600
PUBMED
11591384
Contact: Karen G. Welinder
Instituit for bioteknologi
Aalborg Universitet
COMMENT

```

/cultivar="field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="tuber"
/clone_lib="Mature tuber lambda ZAP"
/note="Vector: Lambda ZAP"

ORIGIN

Query Match 18.0%; Score 167.6; DB 4; Length 182;
Best Local Similarity 95.1%; Pred. No. 3.2e-41; Indels 0; Gaps 0;
Matches 173; Conservative 0; Mismatches 9;
Db 199 CAGGGCTGCGTCGTATGTTTCATGAAGTGCAGGATGTTTCATGACCAACTGTATGCTTC 258
182 CAGGGCTGCGTCGTATGTTTCATGAAGTGCAGGATGTTTCATGACCAACTGTATGCTTC 123
Qy 259 AATAACACCCCATCGGACGTAGCATGCTGTTTCAACTATGCGCAAAATGTT 318
Db 122 AATAACACCCCATCGGACGTAGCATGCTGTTTCAACTATGCGCAAAATGTT 63
Qy 319 CTCGCGGGCTGACAGCCAAATGCTGAAGATACCCAGGTTTGAGCGTCAATCTGTT 378
Db 62 CTCGCGGGCTGACAGCCAAATGCTGAAGATACCCAGGTTTGAGCGTCAATCTGTT 3
Qy 379 AC 380
Db 2 AC 1

RESULT 6

BI406111/c
LOCUS 165 bp mRNA linear EST 11-SEP-2002
DEFINITION 155.0 Mature tuber lambda ZAP Solanum tuberosum cDNA, mRNA
sequence.
ACCESSION BI406111
VERSION BI406111.1 GI:15185525
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 165)
AUTHORS Crookshanks, M., Emmersen, J., Welinder, K.G. and Nielsen, K.L.
TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE 21475600
PUBMED 11591384
COMMENT Contact: Karen G. Welinder
Institute for bioteknologi
Aalborg Universitet
Schigaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk
Sequenced from the 5' end.
High quality sequence stop: 165
POLYA=No.

FEATURES
source

1. .165
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="field grown Kuras"
/db_xref="taxon:4113"
/tissue_type="tuber"
/clone_lib="Mature tuber lambda ZAP"
/note="Vector: Lambda ZAP"

ORIGIN

Query Match 14.7%; Score 136.8; DB 4; Length 165;
Best Local Similarity 95.0%; Pred. No. 1.6e-31;
Matches 152; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 236 TTCATGAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATGGCTGTT 295
Db 165 TTCATGAGCAACTGTATGCTTCAATAACACCCCATCGGACGCTACGCAATGGCTGTT 106
Qy 296 CTTCAACTATGGACAAATGTTTCGCGGGCTGACAGCCAAATGCTGGAAGATACC 355
Db 105 CTTCAACTATGGAC-AAATGTTTCGCGGGCTGACAGCCAAATGCTGGAAGATACC 47
Qy 356 CAGGTTTGAGCGTCAATCTGTTACCGAATTCAGCCCC 395
Db 46 CAGGTTTGAGCGTCAATCTGTTACCGAATTCGGGATCC 7

RESULT 7

AF094919
LOCUS 755 bp DNA linear GSS 29-AUG-2000
DEFINITION AF094919 Salmonella typhimurium LT2, Lambda DASH II Salmonella
typhimurium genomic clone 906-T3, genomic survey sequence.
ACCESSION AF094919.1 GI:4322761
VERSION AF094919
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
1 (bases 1 to 755)
REFERENCE Wong, R.M.Y. and McClelland, M.
AUTHORS End Sequences of Salmonella typhimurium LT2 Lambda DASHII Clones,
Li-Cor
JOURNAL Unpublished (1999)
COMMENT Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.

FEATURES
source

1. 755
/organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="906-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"

ORIGIN

Query Match 14.2%; Score 131.8; DB 8; Length 755;
Best Local Similarity 84.6%; Pred. No. 9e-30;
Matches 148; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
Qy 1 ATGGAACGACTAAACGACATGTCGGTGTTCGCAAGTAGTTGAATTCGCTCTTTTACC 60
Db 581 ATGGAACGACTAAACGACATGTCGGTGTTCGCAAGTAGTTGAATTCGCTCTTTTACC 640
Qy 61 GCGCGCCGACAGAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATCAAACTG 120
Db 641 GCGCGCCGACAGAGCTACAGATGAGCGTTTCGTCATCAGTCAGACGGTATGCGGAACTG 700
Qy 121 GAGATGAGTTGCGAGTAAGCTGTTAAACCGTAGACAGCGAGGCTTGGCTGA 175
Db 701 GAGATGAGTTGCGAGTAAGCTGTTAAACCGTAGACAGCGAGGCTTGGCTGA 755

RESULT 8

BZ341676
LOCUS 703 bp DNA linear GSS 06-NOV-2002
DEFINITION BZ341676.1 WGS-Sbicolorf (JMI.07 adapted methyl filtered) Sorghum
bicolor genomic clone ic78a09 5', genomic survey sequence.
ACCESSION BZ341676
VERSION BZ341676.1 GI:24741474
KEYWORDS GSS.


```

Db      442 CGCATCGTCCGCCAGGAAGACTCCGGACTACTCAGCGGACGGTGGCTTACCAGCA 501
Qy      487 ATGGTGGTGTGGCGCGGAAAGCTATCTCACAAATACGGCATACCGGAAACCCGCC 546
Db      502 ATGATCACCTGTGCTCGCCGCGGATTATCTGAAAGAACATCCGGCACCGCAGACCCGAA 561
Qy      547 GAT 549
Db      562 GAT 564

RESULT 10
CNS01FUZ
LOCUS   1027 bp DNA linear GSS 01-JUN-2001
DEFINITION
Anopheles gambiae GSS T7 end of clone 05P23 of NotreDamel library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION
AL142332
VERSION
AL142332.1 GI:7000450
KEYWORDS
GSS.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE
1 (bases 1 to 1027)
Genoscope.
Direct Submission
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
2 (bases 1 to 1027)
Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
Direct Submission
Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES
source
1..1027
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="05F23"
/clone_lib="NotreDamel"
/note="end : T7"

ORIGIN
Query Match 10.6%; Score 98.2; DB 9; Length 1027;
Best Local Similarity 47.3%; Pred. No. 3.9e-19;
Matches 375; Conservative 5; Mismatches 408; Indels 5; Gaps 3;

Qy      76 CTACAGATCAGCGTTTCGTCCTCATGTCAGACGGTATCAAACTCGAAGATCAGTTGCGAG 135
Db      20 CTGCAATTTGTCAGCTCGCGCGGTAGAAAAGCATCGCTGCTCGAGAGCGGACTGGGG 79

Qy      136 GTAAGCTGTTAAACCGTAGCAGCAGCAGCATTTGGCTGACCGAAGCCGGTAGAATTTAC 195
Db      80 GTACGGCCATTTTCAGGCGCACCCCGCAGCCAGACCTTGACCGGATAACGGCGCCCTTTT 139

Qy      196 TACAGGGCTCGCGTGTGTTATGTTATGAGTGCAGGATGTTTCATGACCACTGATGCC 255
Db      140 TATGAACGGTGCCTGCGTGGTGGAGGAATTCGGGGCGCAGATCGCTGCTTGAACG 199

Qy      256 TTCAATAACACCCCATCGGACGGTACCGCATTTGGCTGTCTTCAACTATGCGACAAAAT 315
Db      200 GGGAAACAGCAGGTACAGCGCGCCCTGCGCGTTGCCATCGCGGTACTGTTTGGCGCGCAG 259

Qy      316 GTTCTCGCGGGCTGACAGCCAAAATGCTGAAAGAAATACCCAGGTTTGAGCGTCAATCTG 375

```

```

Db      260 TGGCTCGCGCGCTGTTGATTAACTTGGCGCAGGAACACCCCGGCTTTRACTTGAATG 319
Qy      376 GTTACCGGAATTCACGCCCCGACCTGATTCGCGAGGCTCTGGATGTGGTATCCGCGTC 435
Db      320 TCGTTTTCAGCGACCGCGTGTGGATCTTGGTGGAGGAGGTTTGATATGGCGGTGGGTAAC 379
Qy      436 GGCGCGTTGAGGATTCAGCCTGTTTCCCGCGCTCTG--GGCGGATGCCCAATGGTGG 493
Db      380 GGCACTGTCAGACAGCAGCATGCTGCTGCGCAGAAARCTGGGGAACACCGCATGTC 439
Qy      494 TGTGCGCGGAAAAAGCTATCTCACAAATACGGCATACCGGAAAAACCCGCGGATTTGA 553
Db      440 TGTGCGCTGCGCGGATTATCTRCAGAACAAAAGCCAAACCGCGTGGTGACTTAC 499
Qy      554 GTAGTCATTATGCTTTGAATA-CACGCTGCGCGCCGACAAATGAATTTG--AACTGATCG 610
Db      500 CCATATATGCGCAATTAATCTCTGCGCGCAGGAGAGTGTACCTTGGCACTGATGG 559
Qy      611 CACCGGAAGGGATCTCGACTGCGCTGATCCCAAGAAAGATTGTGACTAATGATCGA 670
Db      560 ATRACGAAGGGACCTCGCACACCTTTACTCCCGCTTCATCGCTCAATATGGATGATTTGC 619
Qy      671 TGACGCTGGTGGCTGCTGCTGACGGCGGGTCCGGGATCCCTACGTGCGCGCTGATGGG 730
Db      620 AGCCATCTGCGACCGCGCTGCGCGGACACGGGATTCCTGGTCCCTCTGCTGATGG 679
Qy      731 TGATCAACGAGATCAATCTGTTGGGAGCTGGAGATCCTGCTGCGCGTTTACCACTGATG 790
Db      680 CCAATTAAGAAATTCATCAGGGAAGCTCTGTCGCTCTCTGAAACAGGCGCCGATGCC 739
Qy      791 CAGCGCGGTTTATGCTGTTATATACGAAAGATAGCTGCGCTGAGGTACAGGTGCG 850
Db      740 ATTTGACGTTTCATCGCTGCTGCGACAAACCGCACCTGCGCGTGGAGTGAGGATTG 799
Qy      851 TGATCAACTCGCT 863
Db      800 CCGTCGATACGCT 812

RESULT 11
BZ553247 1178 bp DNA linear GSS 17-DEC-2002
pacsl-60_405.si pacsl-60 Pseudomonas aeruginosa genomic clone
pacsl-60_405, genomic survey sequence.
BZ553247
BZ553247.1 GI:27159019
GSS.
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1178)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 35245, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..1178
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacsl-60_405"

FEATURES
source

```

```

/clone_lib="pacsl-60"
/notes="clinical isolate 1-60 whole genomic shotgun
library."

ORIGIN
Query Match      10.2%; Score 95; DB 8; Length 1178;
Best Local Similarity 50.1%; Pred. No. 4.1e-18;
Matches 262; Conservative 0; Mismatches 260; Indels 1; Gaps 1;

QY 1 ATGGACGACTAAAGCATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 139 ATGGATTCGCTGAAGGCAATGCGATCTTCGCCACTGTGTGGACAAGGCTCGATGCA 198
QY 61 GCGCGCGCCAGACAGCTACAGATGAGCGTTTCCTCCATCAGTCAGACAGTATCAAACTG 120
DB 199 GCGCGCGCCAGACAGCTGCGCATGACCCCTTCGGCAGTCAGCAGCAGATCGCAAGCTG 258
QY 121 GAAGATGAGTTCAGGTAAGCTGTAAACCTAGACACAGCAGATTTGGCTGACCGAA 180
DB 259 GAGAGCCGCGCCAGGTCACTTTGTGTCATGCGACACCGCCGCGCTGACCTGACCGAG 318
QY 181 GCGGCTAGAAATTTATACAGGCGCTGCGGTCTGTATGCTTCATGAAGTCAGAGATTTTCAT 240
DB 319 GCGCGCGAGCGCTTCATCGCAGTTGCGCGCAGATGCTGGCGATCGCCGAGGAGCCGAG 378
QY 241 GAGCAACT-GTAGCGCTTCAATPACACCCCGCATCGGACGCTACCGATTTGGCTGTTCTTC 299
DB 379 GCGCGCGCTCGCGGTAAATGCGCGATTCGCGGTGCGGAGTTTCGCGCTGCGCGCGCGGT 438
QY 300 AACTATGGCACAAATGTTCTCGCGGCTGACACCCAAATGCTGAAGAAATACCCAGG 359
DB 439 GGGGTTCTCGGACGCTGATCACCGAGCGCTGAAACCGTTGCTGGAGAACCCGCGCA 498
QY 360 TTGAGCGTCAATCTGGTTACCGGAATTCAGCCCCCGACCTGATTTGCGACGCTTGGGA 419
DB 499 GTTGCGCTGCAACTGTTCTTCAGGACGAGGCGCATGATCTGTCGCGGAGCGCATCGA 558
QY 420 TGTGTTGATCCGCTGCGCGCTTCGAGATTCAGCCTGTTTCCGCGCTGTCGGGCGC 479
DB 559 CTGGCGATCCGCGTGGCAATCTCGCGACTCCAGCCTGTTGGCGCGCCACCTCGGCGGA 618
QY 480 GATGCCAATGTGTGTGTCGCGCGCAAAAGCTATCTCACAAA 522
DB 619 CTGGAGCAGCGTGTCTGCGCGCGCGCGCTATTTTGGCCAA 661

RESULT 12
CG899406/c
LOCUS
DEFINITION
pacsl-60_962.sl pacsl-60 Pasteuria penetrans genomic 5',
genomic survey sequence.
ACCESSION
CG899406
VERSION
CG899406.1 GI:39554967
KEYWORDS
GSS.
SOURCE
Pasteuria penetrans
ORGANISM
Pasteuria penetrans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Pasteuria.
REFERENCE
1 (bases 1 to 761)
Opperman,C.H., Davies,K.G., Sosinski,B.R., Waterman,J. and Burke,M.
Unpublished Data
Unpublished (2003)
Contact: Opperman CH
Center for the Biology of Nematode Parasitism and Nematode
Interactions Unit
North Carolina State University and Rothamsted Research, Ltd.
Box 7233, NCSU, Raleigh, NC 27606, USA
Email: warthog@unity.ncsu.edu
Homology: e-val = 1e-40. Description = PROBABLE TRANSCRIPTION
REGULATOR PROTEIN [Raistonia solanacearum]
gi|17431583|emb|CAD18261.1| PROBABLE TRANSCRIPTION REGULATOR
PROTEIN [Raistonia solanacearum] Homology: e-val = 1e-40.
Description = PROBABLE TRANSCRIPTION REGULATOR PROTEIN [Raistonia
solanacearum] gi|17428746|emb|CAD15431.1| PROBABLE TRANSCRIPTION

```

```

REGULATOR PROTEIN [Raistonia solanacearum] row: f column: 7
Class: shotgun
High quality sequence stop: 761.
Location/Qualifiers
1. 761
/organism="Pasteuria penetrans"
/mol_type="genomic DNA"
/strain="Res147"
/db_xref="taxon:86005"
/clone_lib="Res147.1"

ORIGIN
Query Match      9.7%; Score 90.4; DB 9; Length 761;
Best Local Similarity 49.3%; Pred. No. 1e-16;
Matches 266; Conservative 0; Mismatches 271; Indels 3; Gaps 1;

QY 10 CTAAACGCAATGTCGGTGTTCGCAAGTAGTTGAATTTGGCTCTTTTACC GCGCGCGCC 69
DB 573 CTGGAATTTGAAGTATATTTTGTTCAGTCGTGTAATTCGTTAGTTTACCGCGCGCT 514
QY 70 AGACAGCTACAGATGAGCGTTCGTCATCAGTCAGACGCTATCAAACTTGAAGATGAG 129
DB 513 GACGCTTACAGCTGATCGCCCTGCGGTGAGTTCGTCAGTCCGTAACAAGCTGGAAGCAAG 454
QY 130 TTGAGGTAAGCTGTAAACCTAGCAGCAGCAGATTTGGCTGACCGAGCGGTAGA 189
DB 453 CTTGGCGTTAAGCTATTTGACCGCACCGCAGCTCTGAGTATGACCGTTGAAGGGGA 394
QY 190 ATTACTACAGGCGTCCGCTGCTGATGCTTCAAGTAGTCAGATGTTTCATGAGCAACTG 249
DB 393 GAAATTTATCAGCGGCTAAACGCTACTGAGTATGCCAGCATAT---CATGGCTTCT 337
QY 250 TATGCTTCAATTAACACCCCGATCGGAGCGCTACGATTTGGCTGTTCTTCAACTATGGCA 309
DB 336 TACTCTTCAATCAGGACCTCGTGGCCAGCTCGCATTTGATTTCCACTGGCGCTGGCA 277
QY 310 CAAAATGTTCTCGCGGCTGACAGCAAAATCTGAAAGATACCAAGTATTCAGCGTC 369
DB 276 CACAGTATCTGATCCCAATTAAGCGAGTTTAAGGTTTATCTGAAATGAAATA 217
QY 370 AATCTGTTACCGGAATTCAGCGCCCGACCTGATTCGCGACGCTGATGTTGGTATC 429
DB 216 GTCTTAAACGCTTCAGATCGCTGACCCATTTAATTCAGAGAGTGTGGACTGTGTTATT 157
QY 430 CGCGTCGCGCGTTCAGGATTCAGCGCTTTTCCGCGCTGTCGGCGGATGCCAATG 489
DB 156 CGCCTCGGTGAGCTGCTGATTCGAATTTTATTCGCGCAGAAATGGCAGCTGSCAGATG 97
QY 490 GTGCTGTGCGCGCGCAAAAGCTATCTCACAAATAGCGCATACCGGAAACCGCGCGAT 549
DB 96 GTACCTGTGCGACACCCCTTATCTGGATAATATGCGACTCTCTTAACGCTGATGAT 37

```

```

RESULT 13
BZ558369
LOCUS
DEFINITION
pacsl-60_962.sl pacsl-60 Pasteumonas aeruginosa genomic clone
pacsl-60_962, genomic survey sequence.
ACCESSION
BZ558369
VERSION
BZ558369.1 GI:27172622
KEYWORDS
GSS.
SOURCE
Pasteumonas aeruginosa
ORGANISM
Pasteumonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteumonadales;
Pasteumonadaceae; Pasteumonas.
REFERENCE
1 (bases 1 to 954)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pasteumonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center

```

University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .954
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacs1-60.962"
/clone_lib="pacs1-60"
/note="clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

```

Query Match      9.6%; Score 89; DB 8; Length 954;
Best Local Similarity 49.2%; Pred. No. 38-16;
Matches 283; Conservative 0; Mismatches 270; Indels 2; Gaps 1;

QY 2 TGAACGACTAAACGCGATGTCGGTGTTCGCCAAGTAGTTGAATTTGGCTCTTTTACCG 61
DB 75 TGGATCACTACCGCTCTCAAGGTTCTCCGGCGGTGGCGGAATGGCGCTTCGGCG 134
QY 62 CGCCGCCAGACAGCTACAGATGAGGTTTCGTCCATCAGTCAGTCAGCGGTATCAAACTGG 121
DB 135 CGCGGCCCGGAGATGAATCTCTCCGCCGCGCGGTGAGCAAGAACGTCCCGAGCTGG 194
QY 122 AAGATGAGTTGAGGTAAGAGCTGTTAAACCGTAGCACACGAGCAATTCGCTGACCGGAAG 181
DB 195 AAGCGACCTCAAGGTGCCCATGATCAATCGCACACCGCGAGCATGAGCTGACCGAGG 254
QY 182 CGGTGAGATTTACTACAGAGGCTGCGCTGATGCTTCATGAGTGAAGTGAAGTTCATG 241
DB 255 CGCGGGAAGTCTACCGGAGCGCTTGGAGCGCATCTCGAGCACCTCGAGGCCCGCGAGC 314
QY 242 AGCAACTGATGCTTCAATAACACCCCATCGGAGCGCTAGCATTTGGCTGTTCTTCAA 301
DB 315 CGCGCTCACTTGATGAGAGAGGCGCCAGCGGCTGCTGGGCTAGCGCCCGCTGA 374
QY 302 CTATGGCAAAATGTTCTCGCGGGGTGACAGCCAAATGCTGAAGAAATACCCAGGTT 361
DB 375 CCTCGCCCTCACTGCTGACCGGGCCATTCGGGCTTTTCTCCAGGCTTATCCGAGC 434
QY 362 TGAGGCTCAATCTGTTACCGGAATTCAGCCCCGAGCTGATTCGCCAGCTG--GA 419
DB 435 TCGCCCTGGAATGCTCTCTGAGAGCGGCGCGCAGAGCTGATCCGCCAAGAGTCGACCT 494
QY 420 TGTGTGATCCCGCTCGGCGGTTCGAGGAATTCAGACCTGTTTTTCCCGCTCTGGGCGC 479
DB 495 GSCCTGCGTGTGAGACCGGTTTCGCGACTCCGCGCTGTTGCGCGCGCTGCTGT 554
QY 480 GATGCCAATGTTGTTGCGCGCGGAAGCTTATCTACACATACGACATACCG 534
DB 555 CTTGAGACGACGTGCTCTGCGCGCGCGCGCTTACCTGATCAGCATGCGCCGCG 609

```

RESULT 14

BZ567908 1279 bp DNA linear GSS 17-DEC-2002
LOCUS
DEFINITION
pacs2-164_7276.xl pacs2-164 Pseudomonas aeruginosa genomic clone
BZ567908 7276 bp genomic survey sequence.
ACCESSION
VERSION BZ567908.1 GI:27200060
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 1279)
AUTHORS
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

TITLE Whole-Genome-Sequence variation among multiple isolates of
JOURNAL Pseudomonas aeruginosa library
COMMENT J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

source
1. .1279
Location/Qualifiers
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164.7276"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

```

Query Match      9.4%; Score 87.6; DB 8; Length 1279;
Best Local Similarity 49.0%; Pred. No. 9e-16;
Matches 234; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 1 ATGGAACGACTAAACGCGATGTCGGTGTTCGCCAAGTAGTTGAATTTGGCTCTTTTACC 60
DB 114 ATGGAATTCGCTGAAGGCGATGGCGATCTTCGCCACTGAGGTGCGACAAGGCTCGATGGCA 173
QY 61 GCGCGCGCGAGACAGCTACAGATGAGCGTTTCGTCCATCAGTCAGACGGGTATCAAACTG 120
DB 174 GCGCGCGCGAGAGCTTGGCGATGACCCCTTGGCAGTCAGCCAGATCCGCAAGCTG 233
QY 121 GAAGATGAGTTGAGGTAAGCTGTTAAACCGTAGCACACGAGCATTTGGCTGACCGAA 180
DB 234 GAGAGCGCGCGCGAGGTCACTTTGCTGATCGCACACCCCGCGGCTGACCCCTGACCGAG 293
QY 181 GCGGTTAGATTTACTACAGGCGTCCGCTGATGCTTCATGAGTGCAGATGTTTCAAT 240
DB 294 GCGCGCGAGGCTTCTATCGAGTTGCGCGAGATGCTGGCGATGCCGAGGAAGCCGAG 353
QY 241 GAGCAACTGTATGCTTCAATAACACCCCATCGGAGCGCTACGATTTGGCTGTTTCA 300
DB 354 CGGTGCTTTGGCGAATGGCGCGATGCGCGGTGGGTGAGTTGCGCTGCGCGCACAGCG 413
QY 301 ACTATGGCAAAATGTTCTCGCGGGGTGACAGCCAAATGCTGAAGAAATACCCAGT 360
DB 414 GGGTTCTCCGGCAGCTGATCACCCAGGCGCTGAATTCGTTTGGAGAACCCAGCGACAG 473
QY 361 TTGAGCGTCAATCTGTTTACCGGAATTCAGCCCCCGACCTGATTCGCCAGCGTCTGGAT 420
DB 474 TTGCGCTCGAGCTTGTCTTCCAGGAGAGCGCATCGAGCTGGCGCGCGAGCATCGCA 533
QY 421 GTGGTGAATCGCGTTCGCGCGGTTCAGGATTCAGCGCTGTTTCCCGCGCTTGGCGG 478
DB 534 CTGACGATACGCGTTCGAGCATTTCTCCGACTCCAGCTGTCGCGACCATCTTTGTG 591

```

RESULT 15

CG89239/c
LOCUS
DEFINITION
pastsbac094xd07.bl.abi Res147 1 Pasteuria penetrans genomic 5',
genomic survey sequence.
ACCESSION
VERSION CG89239.1 GI:39554734
KEYWORDS
SOURCE
ORGANISM
Pasteuria penetrans
Bacteria; Firmicutes; Bacillales; Alicyclobacillaceae; Pasteuria.
REFERENCE
1 (bases 1 to 741)

Search completed: December 19, 2004, 02:46:57
Job time : 4053.28 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:49:04 ; Search time 1924.01 Seconds
(without alignments)
9659.452 Million cell updates/sec

Title: US-10-759-889-3

Perfect score: 393
Sequence: 1 tccaacctcaaacgacagt.....ctttaagtcagagtgaata 393

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.btg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	100.0	110000	1	ECOUW67_1
2	393	100.0	110000	1	ECOUW67_33
3	393	100.0	290452	2	AC079167
4	391.4	99.6	292497	1	AE016989
5	389.8	99.2	307962	1	AP002564
6	386.6	98.4	300359	1	AE016767
7	362.8	92.3	115904	1	AE005551
8	289.8	73.7	21252	1	AE008855
9	289.8	73.7	258050	1	AL627278
10	289.8	73.7	299991	1	AE016845
11	226	57.5	11744	1	AE015338
12	225.4	57.4	11734	1	AE015339
13	146.6	37.3	258	6	AR389738
14	125.8	32.0	593	1	AB080601
15	117.4	29.9	10553	1	AE013617
16	117.4	29.9	229338	1	AE013617
17	117.4	29.9	235050	1	AJ414158
18	100.4	25.5	110000	1	ECOUW67_1
19	100	25.4	100	6	AX988997

C 20	100	25.4	100	6	AX988998	Sequence
C 21	100	25.4	100	6	AX988999	Sequence
C 22	87	22.1	11593	1	AE005552	Escherich
C 23	65.8	16.7	20157	1	AY647257	Yersinia
C 24	45.6	11.6	302614	1	AE016778	Pseudomon
C 25	44.4	11.3	1512	1	AB077386	Serratia
C 26	44.2	11.2	7218	6	I66494	Sequence 14
C 27	42.4	10.8	348525	1	BX640428	Bordetell
C 28	41.8	10.6	171655	2	BR950210	Danio rer
C 29	41.8	10.6	193489	5	EX005088	Zebrafish
C 30	41.8	10.6	217967	2	EX323049	Danio rer
C 31	40.8	10.4	348642	1	EX640446	Bordetell
C 32	40.6	10.3	303226	1	AE016774	Pseudomon
C 33	40	10.2	154091	2	CR394557	Danio rer
C 34	38.8	9.9	349174	1	AB063522	Wiggleswo
C 35	38.4	9.8	986	8	AB118012	Oryza sat
C 36	38.4	9.8	1181	6	E14788	CDNA encodi
C 37	38.4	9.8	157205	8	AP005098	Oryza sat
C 38	38.2	9.7	109619	2	AC141815	Continuation (4 of
C 39	37.8	9.6	85234	8	AC139743	Medicago
C 40	37.8	9.6	224562	10	AC123880	Mus muscu
C 41	37.6	9.6	205587	2	AC139761	Mus muscu
C 42	37.2	9.5	1035	6	AR384449	Sequence
C 43	37.2	9.5	50434	2	AC115729	Mus muscu
C 44	37.2	9.5	181337	10	AL929046	Mouse DNA
C 45	37.2	9.5	189589	2	AC146641	Otolemur

ALIGNMENTS

RESULT 1

ECOUW67_1

WPCOMMENT

Sequence split into 4 fragments LOCUS ECOUW67 Accession U18997

Fragment Name	Begin	End
ECOUW67_0	1	110000
ECOUW67_1	100001	210000
ECOUW67_2	200001	310000
ECOUW67_3	300001	372438

Continuation (2 of 4) of ECOUW67 from base 100001 (U18997 Escherichia coli K-12 chromoso

Query Match 100.0%; Score 393; DB 1; Length 110000;

Best Local Similarity 100.0%; Pred. No. 8e-104;

Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTCACTCTCAAGCAAGACAGTCGGGATATCAATAAAGCAATTCGTTGGCAAGTACCGCGCA 120

Db 69872 TTCACTCTCAAGCAAGACAGTCGGGATATCAATAAAGCAATTCGTTGGCAAGTACCGCGCA 69931

Qy 61 TGAACACGCGGATGCCAGACAGTCGTAGATACCTGTTGGCAAGTACCGCGCA 120

Db 69932 TGAACACGCGGATGCCAGACAGTCGTAGATACCTGTTGGCAAGTACCGCGCA 69991

Qy 121 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCGTTGGCAAGTACCGCGCA 180

Db 69992 CCAGCCAGAAAATCGCCAGTGTATAAAGCAATTCGTTGGCAAGTACCGCGCA 70051

Qy 181 CAAACACCAAGTAAACGGGAAACAGACTCATGTGACCTTGGTTGTAACAGAGAGCAGG 240

Db 70052 CAAACACCAAGTAAACGGGAAACAGACTCATGTGACCTTGGTTGTAACAGAGAGCAGG 70111

Qy 241 CGTTATTATTTTCAGCATCTGCGCGCAGAGAGGCGATGGAAGCCGGCGAGAGCAA 300

Db 70112 CGTTATTATTTTCAGCATCTGCGCGCAGAGAGGCGATGGAAGCCGGCGAGAGCAA 70171

Qy 301 CATTGCTGTAGATGATATTTAATATATATATAGCGTAACCTGTTATCTATATATAT 360

Db 70172 CATTGCTGTAGATGATATTTAATATATATATAGCGTAACCTGTTATCTATATATAT 70231

Qy 361 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393

Db 70232 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 70264

RESULT 2
U00096_33
WPCOMMENT

Sequence split into 47 fragments LOCUS U00096 Accession U00096

Fragment Name	Begin	End
U00096_00	1	110000
U00096_01	100001	210000
U00096_02	200001	310000
U00096_03	300001	410000
U00096_04	400001	510000
U00096_05	500001	610000
U00096_06	600001	710000
U00096_07	700001	810000
U00096_08	800001	910000
U00096_09	900001	1010000
U00096_10	1000001	1110000
U00096_11	1100001	1210000
U00096_12	1200001	1310000
U00096_13	1300001	1410000
U00096_14	1400001	1510000
U00096_15	1500001	1610000
U00096_16	1600001	1710000
U00096_17	1700001	1810000
U00096_18	1800001	1910000
U00096_19	1900001	2010000
U00096_20	2000001	2110000
U00096_21	2100001	2210000
U00096_22	2200001	2310000
U00096_23	2300001	2410000
U00096_24	2400001	2510000
U00096_25	2500001	2610000
U00096_26	2600001	2710000
U00096_27	2700001	2810000
U00096_28	2800001	2910000
U00096_29	2900001	3010000
U00096_30	3000001	3110000
U00096_31	3100001	3210000
U00096_32	3200001	3310000
U00096_33	3300001	3410000
U00096_34	3400001	3510000
U00096_35	3500001	3610000
U00096_36	3600001	3710000
U00096_37	3700001	3810000
U00096_38	3800001	3910000
U00096_39	3900001	4010000
U00096_40	4000001	4110000
U00096_41	4100001	4210000
U00096_42	4200001	4310000
U00096_43	4300001	4410000
U00096_44	4400001	4510000
U00096_45	4500001	4610000
U00096_46	4600001	4639675

Continuation (34 of 47) of U00096 from base 3300001 (U00096 Escherichia coli K-12 MG1655)

Query Match	Best Local Similarity	100.0%; Score 393; DB 1; Length 110000;
Matches 393; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY 1	TTCAACCTCAAGCAAGTCGCGATATCAATAAAGCAAGCAAGCAAGTACGCGGGTGT 60	
DB 87149	TTCAACCTCAAGCAAGTCGCGATATCAATAAAGCAAGCAAGCAAGTACGCGGGTGT 87208	
QY 61	TGAACAACGCGGATCCAGACAAAGTCGTAGATCTGTGGCAAGTACCCGCGCA 120	
DB 87209	TGAACAACGCGGATCCAGACAAAGTCGTAGATCTGTGGCAAGTACCCGCGCA 87268	
QY 121	CCAGCCAGAAATCGCGAGTGAATAAGCAATTCAAAAATATCGTGGAGGACGCC 180	
DB 87269	CCAGCCAGAAATCGCGAGTGAATAAGCAATTCAAAAATATCGTGGAGGACGCC 87328	
QY 181	CAACACACCGATAAACGGGAACAGACTCATCTGTGACCTTGGTTGTAAGAGAGACGAG 240	

Db 87329	CAACACACCGATAAACGGGAACAGACTCATCTGTGACCTTGGTTGTAAGAGAGACGAG 87388	
QY 241	CGTTATTATTTTCAGCATCTGTGCGCGAGAGAGGCGATGAAAGCGGGCGAGAGCAA 300	
Db 87389	CGTTATTATTTTCAGCATCTGTGCGCGAGAGAGGCGATGAAAGCGGGCGAGAGCAA 87448	
QY 301	CATTGCTGTAGATTGATATTATATATATTAGCGTACTGTTATGCTGTTATCTATATTAT 360	
Db 87449	CATTGCTGTAGATTGATATTATATATTAGCGTACTGTTATGCTGTTATCTATATTAT 87508	
QY 361	GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393	
Db 87509	GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 87541	

RESULT 3
AC079167/c

LOCUS AC079167 290452 bp DNA linear HTG 15-MAY-2002
DEFINITION Mus musculus chromosome 6 clone RP23-226D23 strain C57BL6/J, ***
SEQUENCING IN PROGRESS ***, 196 unordered pieces.

ACCESSION AC079167.2 GI:9945047

VERSION HTG; HTGS PHASE1.

KEYWORDS Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 290452)

Grillis, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,

Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,

Gordon, M., Goltz, J.S. and Kucherlapati, R.

High Throughput Mouse Sequencing

Unpublished

2 (bases 1 to 290452)

Grillis, G., Han, J., Montgomery, K.T., Lee, E., Long, J., Pomerantz, R.,

Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,

Gordon, M., Goltz, J.S. and Kucherlapati, R.

Direct Submission

Submitted (23-AUG-2000) Department of Molecular Genetics, Albert

Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,

Bronx, NY 10463, USA

On Aug 30, 2000, this sequence version replaced gi:9885883.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpcgg.org/Sequence/mouse.html>

Contact: hpcg@mendel.mgh.harvard.edu

-----Summary Statistics

Center project name: AEN

Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 196511 at least Q20

*Consensus quality: 172308 at least Q30

*Consensus quality: 138790 at least Q40

Estimated insert size: agarose-FP - N/A

**Estimated insert size: 286552 sum-of-contigs

Quality coverage: agarose-FP - N/A

Quality coverage: 3.3 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 196 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 852: contig of 852 bp in length

* 83 872: gap of unknown length

* 873 1753: contig of 881 bp in length

* 1754 1773: gap of unknown length

* 1774 2658: contig of 885 bp in length

* 2659 2678: gap of unknown length
* 2679 3514: contig of 836 bp in length
* 3515 3534: gap of unknown length
* 3535 4449: contig of 915 bp in length
* 4450 5384: gap of unknown length
* 5385 5404: contig of 915 bp in length
* 5405 6273: contig of 869 bp in length
* 6274 6293: gap of unknown length
* 6294 7095: contig of 802 bp in length
* 7096 7115: gap of unknown length
* 7116 7981: contig of 866 bp in length
* 7982 8002: gap of unknown length
* 8003 8749: contig of 728 bp in length
* 8750 9875: gap of unknown length
* 9876 9895: contig of 1126 bp in length
* 9896 10631: gap of unknown length
* 10632 10651: contig of 736 bp in length
* 10652 11577: gap of unknown length
* 11578 11597: contig of 926 bp in length
* 11598 12454: gap of unknown length
* 12455 12474: contig of 857 bp in length
* 12475 13344: gap of unknown length
* 13345 13354: contig of 860 bp in length
* 13355 14279: gap of unknown length
* 14280 15245: contig of 925 bp in length
* 15246 15265: gap of unknown length
* 15266 16058: contig of 793 bp in length
* 16059 16078: gap of unknown length
* 16079 16945: contig of 867 bp in length
* 16946 16965: gap of unknown length
* 16966 17850: contig of 885 bp in length
* 17851 17870: gap of unknown length
* 17871 18722: contig of 852 bp in length
* 18723 18742: gap of unknown length
* 18743 19556: contig of 814 bp in length
* 19557 19576: gap of unknown length
* 19577 20816: contig of 1240 bp in length
* 20817 20836: gap of unknown length
* 20837 21756: contig of 920 bp in length
* 21757 21776: gap of unknown length
* 21777 22517: contig of 741 bp in length
* 22518 22537: gap of unknown length
* 22538 23457: contig of 920 bp in length
* 23458 23477: gap of unknown length
* 23478 24289: contig of 812 bp in length
* 24290 24309: gap of unknown length
* 24310 25249: contig of 940 bp in length
* 25250 25689: gap of unknown length
* 25690 26135: contig of 866 bp in length
* 26136 26155: gap of unknown length
* 26156 27048: contig of 893 bp in length
* 27049 28003: contig of 935 bp in length
* 28004 28023: gap of unknown length
* 28024 28874: contig of 851 bp in length
* 28875 28894: gap of unknown length
* 28895 29777: contig of 883 bp in length
* 29778 29797: gap of unknown length
* 29798 30160: contig of 363 bp in length
* 30161 30180: gap of unknown length
* 30181 31014: contig of 834 bp in length
* 31015 31034: gap of unknown length
* 31035 31880: contig of 846 bp in length
* 31881 31900: gap of unknown length
* 31901 33160: contig of 1260 bp in length
* 33161 33180: gap of unknown length
* 33181 34033: contig of 853 bp in length
* 34034 34053: gap of unknown length
* 34054 34572: contig of 919 bp in length
* 34573 34992: gap of unknown length
* 34993 35875: contig of 883 bp in length
* 35876 35895: gap of unknown length
* 35896 36731: contig of 836 bp in length
* 36732 36751: gap of unknown length
* 36752 37639: contig of 888 bp in length
* 37640 37659: gap of unknown length
* 37660 38499: contig of 840 bp in length
* 38500 38519: gap of unknown length
* 38520 39348: contig of 829 bp in length
* 39349 39368: gap of unknown length
* 39369 40377: contig of 1009 bp in length
* 40378 40397: gap of unknown length
* 40398 41469: contig of 1072 bp in length
* 41470 41489: gap of unknown length
* 41490 41890: contig of 401 bp in length
* 41891 42795: contig of 885 bp in length
* 42796 42815: gap of unknown length
* 42816 43744: contig of 929 bp in length
* 43745 43764: gap of unknown length
* 43765 44628: contig of 864 bp in length
* 44629 44648: gap of unknown length
* 44649 45958: contig of 1310 bp in length
* 45959 45978: gap of unknown length
* 45979 46827: contig of 849 bp in length
* 46828 46847: gap of unknown length
* 46848 47807: contig of 860 bp in length
* 47808 47827: gap of unknown length
* 47828 48376: contig of 549 bp in length
* 48377 48396: gap of unknown length
* 48397 49297: contig of 901 bp in length
* 49298 49317: gap of unknown length
* 49318 50182: contig of 865 bp in length
* 50183 50202: gap of unknown length
* 50203 51054: contig of 852 bp in length
* 51055 51074: gap of unknown length
* 51075 51373: contig of 299 bp in length
* 51374 51393: gap of unknown length
* 51394 52053: contig of 660 bp in length
* 52054 52073: gap of unknown length
* 52074 52944: contig of 871 bp in length
* 52945 52964: gap of unknown length
* 52965 53539: contig of 575 bp in length
* 53540 53559: gap of unknown length
* 53560 54374: contig of 815 bp in length
* 54375 54394: gap of unknown length
* 54395 55253: contig of 859 bp in length
* 55254 55273: gap of unknown length
* 55274 56527: contig of 1254 bp in length
* 56528 56547: gap of unknown length
* 56548 57376: contig of 829 bp in length
* 57377 57396: gap of unknown length
* 57397 57473: contig of 77 bp in length
* 57474 58337: gap of unknown length
* 58338 58357: gap of unknown length
* 58358 58444: contig of 87 bp in length
* 58445 58464: gap of unknown length
* 58465 59288: contig of 824 bp in length
* 59289 59308: gap of unknown length
* 59309 60499: contig of 1191 bp in length
* 60500 60519: gap of unknown length
* 60520 61859: contig of 1340 bp in length
* 61860 61879: gap of unknown length
* 61880 62752: contig of 873 bp in length
* 62753 62772: gap of unknown length
* 62773 63623: contig of 851 bp in length
* 63624 63643: gap of unknown length
* 63644 64714: contig of 1051 bp in length
* 64715 65167: contig of 453 bp in length
* 65168 65187: gap of unknown length

Query Match	100.0%	Score 393	DB 2	Length 290452
Best Local Similarity	100.0%	pred. No. 8.6e-104		
Matches 393	Conservative 0	Mismatches 0	Indels 0	Gaps 0
1	TTCAACCTCAAACGAAAGTCCGATATCAATTAACAAACGACGCAATAGAGCGCGGTGT	60		
53527	TTCAACCTCAAACGAAAGTCCGATATCAATTAACAAACGACGCAATAGAGCGCGGTGT	53468		
61	TGAACAACGCCCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCA	120		
53467	TGAACAACGCCCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCA	53408		
121	CCAGCCAGAAATCGCCAGTGAATAAGCAATTCAAAAATATCGTGGAGAGACAGCC	180		
53407	CCAGCCAGAAATCGCCAGTGAATAAGCAATTCAAAAATATCGTGGAGAGACAGCC	53348		
181	CAACACACAGATACCGGAAACAGACATCTGTTGACCTCGTTGTAAAGAGAGAGCAGG	240		
53347	CAACACACAGATACCGGAAACAGACATCTGTTGACCTCGTTGTAAAGAGAGAGCAGG	53288		
241	CGTTATTATTTTCAGCATCTGTGCGCGCAGAGAGGGCATCGAAAGCGCGCGAGAGCAA	300		
53287	CGTTATTATTTTCAGCATCTGTGCGCGCAGAGAGGGCATCGAAAGCGCGCGAGAGCAA	53228		
301	CATTGCTGTAGATTCATTTAATATATATCGGTAACTGTTATCTGTTATCTATATAT	360		
53227	CATTGCTGTAGATTCATTTAATATATATCGGTAACTGTTATCTGTTATCTATATAT	53168		
361	GTGATCTAAATCACCTTTAAAGTCAGAGTGAATA	393		
53167	GTGATCTAAATCACCTTTAAAGTCAGAGTGAATA	53135		

RESULT 4	292497 bp	DNA	linear	BCT 22-APR-2003
AE016989	Shigella flexneri 2a str. 2457T			
LOCUS	Shigella flexneri 2a str. 2457T			
DEFINITION	Shigella flexneri 2a str. 2457T section 12 of 16 of the complete genome			
ACCESSION	AE016989			
VERSION	AE016989.1			
KEYWORDS	GI:30042681			
SOURCE	Shigella flexneri 2a str. 2457T			
ORGANISM	Shigella flexneri 2a str. 2457T			
REFERENCE	Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A., Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S., Schwartz, D.C. and Blattner, F.R.			
AUTHORS	Complete Genome Sequence and Comparative Genomics of Shigella flexneri Serotype 2a Strain 2457T			
JOURNAL	Infect. Immun. 71 (5), 2775-2786 (2003)			
PUBLISHED	12704152			
REFERENCE	2 (bases 1 to 292497)			
AUTHORS	Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G. III, Rose, D.J., Darling, A., Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S., Schwartz, D.C. and Blattner, F.R.			
TITLE	Direct Sequencing			
JOURNAL	Submitted (13-JUN-2002) Genetics Laboratory, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA			
FEATURES	Location/Qualifiers			
source	1..292497			
gene	/organism="Shigella flexneri 2a str. 2457T"			
	/mol_type="genomic DNA"			
	/strain="2457T"			
	/serotype="2a"			
	/db_xref="taxon:198215"			
	98..1169			
	/gene="vgjN"			
	/locus_tag="S3326"			
	/note="frameshift"			

```

/translation="MTAIAIDLOAGEKLTAVAPFLAGIQNEBQYTOALELVDHLLND
PENPLDLVCAKITAWESAPFAEFNAQAPGCGIAVIRITMDQYGLTISDLPEIG
SKSMVSRVLSGKELTLEHAKLATRGISPALFID"
complement(6239..6553)
/gene="ygjN"
/locus_tag="S3330"
complement(6239..6553)
/gene="ygjN"
/locus_tag="S3330"
/notes="residues 1 to 104 of 104 are 100.00 pct identical
to residues 1 to 104 of 104 from Escherichia coli K-12 :
B3083"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18408.1"
/db_xref="GI:30042685"
/translation="WHLITOKALKDAAEKYPOKHTELVALGNITAKYFKKPKSLKAV
FPLDNFKYLDKHYFNVCGNLRVAVVFFSQKCYREVMTHKEYDFFAVHTKG
K"
6736..6822
/notes="REP (repetitive extragenic palindromic) element,
contains 2 REP sequences"
complement(6838..7974)
/gene="ygjO"
/locus_tag="S3331"
complement(6838..7974)
/gene="ygjO"
/locus_tag="S3331"
/function="putative enzyme; Not classified"
/notes="residues 1 to 378 of 378 are 95.76 pct identical to
residues 11 to 388 of 388 from Escherichia coli K-12 :
B3084"
/codon_start=1
/transl_table=11
/product="putative enzyme"
/protein_id="AAP18409.1"
/db_xref="GI:30042686"
/translation="MSHLNDGFRSLTLOFPAFDVNPLOAWEADEVILLQOLDDEI
LGPVLLIDAGALSALAEKHPYSIGSYISELATRENRLKNGIDESSVFLDSTAD
YPOQGVLLIKFKTLNLEQOALRKRVVSDTRIITAGAKARDIHSTLFLFKVJG
PPTTLAWKRLNLTNFBPLADAPQVSKLEGTDWTIHNHANVFSRTGLDIGAR
FMQHLNLEGEIVDLGCGNGVI GLTLDDKNPQAKVVFDESPMAVASSRLNVETNM
PEALDRSEFMNNALSGVEPRFNVLNCPFPFOQHATLDNVAMFWHYARRCKLING
ELYIVANRHLDYFHKLKKIFGNCCTTIAINNKFVVLKAVLGRRR"
8023..8562
/gene="ygjP"
/locus_tag="S3332"
8023..8562
/gene="ygjP"
/locus_tag="S3332"
/notes="residues 1 to 179 of 179 are 98.32 pct identical to
residues 1 to 179 from Escherichia coli K-12 :
B3085"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18410.1"
/db_xref="GI:30042687"
/translation="MSJPFPSNVMNLTLYQYPEQLLSQVRLTINQRIGDVLA
KRYFGTHYATDKALQYTDQKQFLNAPPKMVDNKHVKNALGHTAVSRV
QCSKLKAAEIRVATVFNPAEPFLRMIVVHELAKHKEKHNKAFYQLCCHMEPQYHQ
LEFDTRLWLTQLSLGQDKI"
8639..9331
/gene="ygjQ"
/locus_tag="S3333"
8639..9331
/gene="ygjQ"
/locus_tag="S3333"
/notes="residues 1 to 230 of 230 are 97.82 pct identical to
residues 1 to 230 from Escherichia coli K-12 :
B3086"
/codon_start=1

```

```

/transl_table=11
/product="hypothetical protein"
/protein_id="AAP18411.1"
/db_xref="GI:30042688"
/translation="MLRAFARLLLRICFSRRTUKIACILLLVAGATIFIADRVVMNAS
KQLTWGDVNAVPARNVGLLGLGAPRGNYFTRRIDTAAALYHAGKVWLLVSGDNGRKN
YDEASGMQOALIAKGVPAKVI FCDYAGFSTLDSVVRKAKYFGENHITIIISQEFHNORT
IWLAKOYGIIDAIGFNAPDLNMKNHGFYTQLREKLARVSAVIDAKILHEQPKYLGPSVMI
GPRSEHSCPAKE"
9392..10408
/gene="ygjR"
/locus_tag="S3334"
9392..10408
/gene="ygjR"
/locus_tag="S3334"
/notes="residues 1 to 334 of 338 are 98.80 pct identical to

```

```

Query Match 99.6%; Score 391.4; DB 1; Length 292497;
Best Local Similarity 99.7%; Pred. No. 2.5e-103;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTCAACCTCAAACGACAGTCGCGATATCAATAAACAAGCAGCATAGACGCGGTGT 60
DB 160258 TTCACCTCAAACGACAGTCGCGATATCAATAAACAAGCAGCATAGACGCGGTGT 160317

QY 61 TGAACAAACGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAAGTACCGGCGCA 120
DB 160318 TGAACAAACGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAAGTACCGGCGCA 160377

QY 121 CCAGCCAGAAATCGCCAGTGATAAAGCAATTCMAAATATATCGTGGGAAGACAGCC 180
DB 160378 CCAGCCAGAAATCGCCAGTGATAAAGCAATTCMAAATATATCGTGGGAAGACAGCC 160437

QY 181 CAAACACACGATAACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGCAGG 240
DB 160438 CAAACACACGATAACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGAGAGAGCAGG 160497

QY 241 CGTTATTTTTCAGCATCTGTCCGCGCAGAGAAGGCGATGGAAGCCGGCGGAGACAA 300
DB 160498 CGTTATTTTTCAGCATCTGTCCGCGCAGAGAAGGCGATGGAAGCCGGCGGAGACAA 160557

QY 301 CATTGCTGTAGTTGATTTAATATATATTAGCGTAACTGTTGCTGTGTTATCTATTAT 360
DB 160558 CATTGCTGTAGTTGATTTAATATATATTAGCGTAACTGTTGCTGTGTTATCTATTAT 160617

QY 361 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 393
DB 160618 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 160650

```

RESULT 5

AP002564

LOCUS

DEFINITION

AP002564 BAO00007

ACCESSION

AP002564.1

VERSION

GI:13363382

KEYWORDS

SOURCE

ORGANISM

Escherichia coli O157:H7

Escherichia coli O157:H7

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

REFERENCE

AUTHORS

Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S.,

Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,

Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,

Sasakawa, C., and Shinagawa, H.

Complete nucleotide sequence of the prophage VT2-Sakai carrying the

verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7

derived from the Sakai outbreak

Genes Genet. Syst. 74 (5), 227-239 (1999)

MEDLINE

20198780

PUBMED

10734605

REFERENCE

2

ARNGLADPGLIGVKGCSVALVLMIFQFPMGLGMFRPVAGLAGGLVALIVLIFCAEII
SRPFRVLIGIGVSWFAAGIGVPMVTDVDTALMWLSGSHHANMVLGISAQWM
LPAALLLRTADIALGHQVATGLGNRSALLRVAAPILITAVCSVGNIGF
VGLIAHISRFRILRGQTHLLGSVAGALLVLLADISGLRALFLPLQUPAGIISLIG
GPFLLLLWQRNSF"
/complement (3931..4914)
/gene="ECs3915"
/complement (3931..4914)
/gene="ECs3915"
/note="probable ferrichrome ABC transporter
(permease), similar to ferrichrome ABC transporters
(permease) e.g. [Synecocystis sp.] gi|7442493|pir|S74438
percent identity 43 in 315 aa. [Bacillus subtilis]
gi|1706795|sp|P49936|PHUB_BACSU percent identity 39"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative permease of ferrichrome ABC
transporter"
/protein_id="BAB37338.1"
/db_xref="GI:13363388"
/translation="WMRTGLVTVIFLALLVGCVVYVPGIGARPTAPOTVLOAFLHPEP
QNFENHIVLRPLRLAALLGSLVAGALLQAVIRNPLGEPIHLGLNAGALVAV
RASALGLAPVRPLASTGGALLFILLLSAGRSGLTPMKVTLGVALSAFVSSI
TRAILLDQTLAMRTWLAGDAGDWTATGSFSLGGVTLAIYLAPELNNMLAG
DRMAGLGVSLRTFTTLLATALLCGAAVSIAGPIGFVGLVPOIVRLWSADLVL
LPSCVAGALLLADIARTLTTPHELATGWTALVAGPVFVIMATRMFX"
/complement (4911..5720)
/gene="ECs3916"
/complement (4911..5720)
/gene="ECs3916"
/note="probable ABC-type iron-siderophore transport system
ATP-binding protein, similar to ABC-type iron-siderophore

gene

CDS

Query Match

Best Local Similarity 99.2%; Score 389.8; DB 1; Length 307962;

Matches 391; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTCAACCTCAACGACAGTCGCGATATCAATAAATAAAGCAGCAATAGACGCGGTGT 60
DB 213062 TTCAACCTCAACGACAGTCGCGATATCAATAAATAAAGCAGCAATAGACGCGGTGT 213121
QY 61 TGAACAAACCGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAGTACCGCGGCA 120
DB 213122 TGAACAAACCGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAGTACCGCGGCA 213181
QY 121 CCAGCCAGAAAATCCCGAGTATGATAAAGCAATTCAAAAATATCGTGGGAGGACAGCC 180
DB 213182 CCAGCCAGAAAATCCCGAGTATGATAAAGCAATTCAAAAATATCGTGGGAGGACAGCC 213241
QY 181 CAACACCAAGTAACCGGAAACAGACATCTGTTGACCTGTTGTTAAAGAGAGAGCAGG 240
DB 213242 CAACACCAAGTAACCGGAAACAGACATCTGTTGACCTGTTGTTAAAGAGAGAGCAGG 213301
QY 241 CGTTATTATTTTCAGCATCTGTCGCCGAGAGAGGCGATGAAAGCGGCGGAGAGCAA 300
DB 213302 CGTTATTATTTTCAGCATCTGTCGCCGAGAGAGGCGATGAAAGCGGCGGAGAGCAA 213361
QY 301 CATTCCTAGATTGATATTAATATATAGCAACTGTTATCTGTTATCTATATAT 360
DB 213362 CATTCCTAGATTGATATTAATATATAGCAACTGTTATCTGTTATCTATATAT 213421
QY 361 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
DB 213422 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 213454

RESULT 6

AE016767

LOCUS

DEFINITION Escherichia coli CF073 section 13 of 18 of the complete genome. BCT 24-DEC-2002

ACCESSION

AE016767.1

VERSION

GI:26110054

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

Escherichia coli CF073
Escherichia coli CF073
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 300359)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Extensive Mosaic Structure Revealed by the Complete Genome Sequence
of Uropathogenic Escherichia coli
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)
12471157
2 (bases 1 to 300359)
Welch,R.A., Burland,V., Plunkett,G.D. III, Redford,P., Roesch,P.,
Rasko,D.A., Buckles,E.L., Liou,S.-R., Boutin,A., Hackett,J.,
Stroud,D., Mayhew,G.F., Rose,D.J., Zhou,S., Schwartz,D.C.,
Perna,N.T., Mobley,H.L.T., Donnenberg,M.S. and Blattner,F.R.
Direct Submission
Submitted (20-JUN-2002) Genetics Laboratory, University of
Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. 300359
/organism="Escherichia coli CF073"
/mol_type="genomic DNA"
/strain="CF073"
/db_xref="taxon:199310"
188..547
/gene="yqiC"
/locus_tag="c3796"
188..547
/gene="yqiC"
/locus_tag="c3796"
/note="Escherichia coli K-12 ortholog: b3042; Escherichia
coli O157:H7 ortholog: z4400"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiC"
/protein_id="AA082241.1"
/db_xref="GI:26110055"
/translation="WTYPTTMAVNTYQVRETTMDPKIKIQIAQVHESMPKQIRE
FGEDVEKKIQTLQAQLTRLDLVSREEDFQVTLTREKLALQRISELEARYNP
VABQSPPAIPVVDKAE"
complement (582..788)
/gene="glsG"
/locus_tag="c3797"
complement (582..788)
/gene="glsG"
/locus_tag="c3797"
/function="putative enzyme; Macromolecule synthesis,
modification: Polysaccharides - (cytoplasmic)"
/note="Escherichia coli K-12 ortholog: b3049; Escherichia
coli O157:H7 ortholog: z4401"
/codon_start=1
/transl_table=11
/product="Glycogen synthesis protein glsG"
/protein_id="AA082242.1"
/db_xref="GI:26110056"
/translation="MLMDHSLNANFDLARSFARMHAEGRPVILAVTGNWDEHR
TWFCARYAYCQMQMTRLELEH"
1052..1681
/gene="yqiJ"
/locus_tag="c3798"
1052..1681
/gene="yqiJ"
/locus_tag="c3798"
/function="putative"
/note="Escherichia coli K-12 ortholog: b3050; Escherichia
coli O157:H7 ortholog: z4402"
/codon_start=1
/transl_table=11
/product="Hypothetical protein yqiJ"

Db 182461 CGTTATTATTTTCAGCATCTGTGCCGCGAGAGGCGCATGGAAAGCGCGGAGAGCAA 182520

Qy 301 CATTCGCTGATGATGATTAATATATATAGCTACTGTTATGCTGTTATCTATATAT 360

Db 182521 CATTCGCTGATGATGATTAATATATATTAACGTAACTGTTATGCTGTTATCTATATAT 182580

Qy 361 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 393

Db 182581 GTGATCTAAATCACTTTTAAAGTCAGAGTGAATA 182613

RESULT 7

AE005551

LOCUS

DEFINITION Escherichia coli O157:H7 EDL933 DNA linear BCT 21-MAR-2001 of 290.

ACCESSION AE005551

VERSION AE005551.1

KEYWORDS GI:12517850

SOURCE

ORGANISM

REFERENCE

AUTHORS

PERNA,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,I., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

PERNA,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D., Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A., Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,I., Grobeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K., Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C., Welch,R.A. and Blattner,F.R.

TITLE

JOURNAL

Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

FEATURES

source

1. .11904

/organism="Escherichia coli O157:H7 EDL933"

/mol_type="genomic DNA"

/strain="EDL933"

/serotype="O157:H7"

/db_xref="taxon:155864"

/note="enterohemorrhagic"

/complement(228..620)

/gene="rpsI"

/note="synonym: Z4588"

/complement(228..620)

/gene="rpsI"

/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"

/note="Residues 1 to 130 of 130 are 100.00 pct identical to residues 1 to 130 of 130 from Escherichia coli K-12 Strain MG1655: B3230"

/codon_start=1

/transl_table=11

/product="30S ribosomal subunit protein S9"

/protein_id="AAG58358.1"

/db_xref="GI:12517851"

/translation="MAENQYGTGRKSSAARVFKPGNKIVINRSLQYFGRETA RMVVRQPLEVDVMEKLDLYITVKGGISGQAGAIRGHIPTALMEYDESURSELKAG FVTRDARQVEKVKGLRKARPPQFSKR"

/complement(636..1064)

/gene="rplM"

CDS

gene

/note="synonym: Z4589"

/complement(636..1064)

/gene="rplM"

/function="structural component; Macromolecule synthesis, modification: Ribosomal proteins - synthesis, modification: Ribosome"

/note="Residues 1 to 142 of 142 are 100.00 pct identical to residues 1 to 142 of 142 from Escherichia coli K-12 Strain MG1655: B3231"

/codon_start=1

/transl_table=11

/product="50S ribosomal subunit protein L13"

/protein_id="AAG58359.1"

/db_xref="GI:12517852"

/translation="MKTIKAPETVKEDVYVDATGKTGLRGLATLARRLRGKHAEY TPVHTGDYIIVLNADKAVTGNKRTDKVYHHTGIGGIKQATFEEMIAARPERVIE IAVKMPLPGKLGAMFRKLKVVAGNEHHAQOQVLDI"

/complement(1283..2410)

/gene="yhcm"

/note="synonym: Z4591"

/complement(1283..2410)

/gene="yhcm"

/function="orf; Unknown function"

/note="Residues 1 to 375 of 375 are 100.00 pct identical to residues 1 to 375 of 375 from Escherichia coli K-12 Strain MG1655: B3232"

/codon_start=1

/transl_table=11

/protein_id="AAG58360.1"

/db_xref="GI:12517853"

/translation="MQSVTQSVLKALNEGSHQPDVQKEAVSRLEIIYOELINSTP PAPRTSGLMARVGLMGKREDTKTPVRGLYMGVGGRGTWMDLDFYQSLPGRKOR LHFHFMRLVHEBLTALOGOTDPLEIIADRFKAETDVLCEDFEVDITAMLLGLLM KALFARGITLAVATSNIPDDELRYNGLRARFLDAIKQHCDDMVNDVAGVDFRLTL TQHLWLSPLHDETQAKMDKLALGAKGKNSPTLEINHRPLATGVENQTLAVST LCVDKASQHDYIALSELTFLVLMFDVPMVTRLMESEARFIALVDFEYRHHVKLVS AEVPLVEIYQCDRLKFEQRCLSRLQEMOSEYLRKREHLAG"

/2598..3002

/gene="yhcb"

/note="synonym: Z4592"

/2598..3002

/gene="yhcb"

/function="orf; Unknown function"

/note="Residues 1 to 134 of 134 are 100.00 pct identical to residues 1 to 134 of 134 from Escherichia coli K-12 Strain MG1655: B3233"

/codon_start=1

/transl_table=11

/product="orf, hypothetical protein"

/protein_id="AAG58361.1"

/db_xref="GI:12517854"

/translation="MEWTWEYALIGLVGIIIGAVAMRFGNRKLRQOQALQVLEKNK AELDEVRELVSHFARSALLDTMAHDYRLQYCHMAKSSSSLLPELSAEPNRRLA ESEASNDQAPVQMPRDYSEASGLLTGAKRD"

/3156..4523

/gene="degQ"

/note="synonym: Z4593"

/3156..4523

/gene="degQ"

/function="enzyme; Macromolecule degradation: Degradation of proteins, peptides, glyco"

/note="Residues 1 to 455 of 455 are 99.78 pct identical to residues 1 to 455 of 455 from Escherichia coli K-12 Strain MG1655: B3234"

/codon_start=1

/transl_table=11

/product="serine endoprotease"

/protein_id="AAG58362.1"

/db_xref="GI:12517855"

/translation="MKKQTLISALALSVLTLTSFQVAVSIPOQVADQAPLPSIAP MLEKVLPAVSVRVEGTASQKQPIPEEFKKFFGDDLPDQAPQFPKGLSGSVIINASKG YVLTNNHVINQAKISIQINDGREFDKIDGSDIALLQIQNPSKLTQITAIADSD KLRVGDFAVAVGNPFGLTGATATSGIVSALGRSLNLEGLFIQTDAINRNGSGAL

gene

LNMLGELICINTAILAPGGSGVIGFPAIPSNMARTLAQQLIDFCEIKRGLIGKTEM
SADIARFNLDVQRGAFVSEVLPSSGSAKAGVXTGDIITSLNGRPLNSPHELRSRIAT
TEPGTKVLKGLKPLEVEVITLDTSTSSASAEIMTPALEGATLSDGQKDGKGI
KIDEVYKGSPPAAQAGLQKDDVIIIVNRDRVNSIAEMRKRVLAAKPAIIALQIVRGNESI
YLLMR"
4613. 5680
/gene="degs"
/note="synonym: Z4594"
4613. 5680
/gene="degs"
/function="enzyme; Macromolecule degradation: Degradation
of proteins, peptides, glyco"
/note="Residues 1 to 355 of 355 are 99.71 pct identical to
residues 1 to 355 of 355 from Escherichia coli K-12 Strain
MG1655: B3235"
/codon_start=1
/transl_table=11
/product="protease"
/protein_id="AAG58363.1"
/db_xref="GI:12517856"
/translation="MFVKLLRSVAIGLIVGAILLVAMPSLRSLNPLSTPQFDSTDETP
ASYNLAARPAAPVNVNIRGLNTHSHQLBIERTLGSVINDQSGYIITNKHVINDAD
QIIVALQDQRFVALLVSDSLDLDVILKINATGGLPTIPINARVPHIGVILAIEN
PYNKQITIGIISATIGLNPTRQNFQDASINHGSGALVSLGELMGINIL
SPDKNDGTPGIGFAIPFQATKIMDKLRIDGRVIRGYIGIGREIAPLHAQGGGI
DOLQGIIVNEVSPDGPANAGIQVNDLIIISVDNKPASIALETMDQVABIRPGSVIPV
VNRDQQLTQVITQIEPATN"
complement(5744..5682)
/gene="mdh"
/note="synonym: Z4595"
complement(5744..5682)
/gene="mdh"
/function="enzyme; Energy metabolism, carbon: TCA cycle"
/note="Residues 1 to 312 of 312 are 100.00 pct identical
to residues 1 to 312 of 312 from Escherichia coli K-12
Strain MG1655: B3235"
/codon_start=1
/transl_table=11
/product="malate dehydrogenase"
/protein_id="AAG58364.1"
/db_xref="GI:12517857"
/translation="MKVAVLGAAGIGQALALLKTOLPSGSELSLYDIAVPTPGVAV
DUSHPTAKVIGFSGEDATPALEGADVILISAGVAPKPMRSDLEFNVAGIKNLV
QQVAKTCPCACIGIINTPVNTTVAIAEVLKAGVYDKNKGFTVILDIIRSNFVAE
LKQKQGEVEVPIVIGSHGVITPLLSQVPGVFTQEVADLTKRIQAGTEVBEAKA
GGGSAITLSQAAARFGLSLVALQEQGVCEYVGDGVQVAFPSQPLLKNGVGE
ERKSICITLSAFQNALQEGMLDLTKKDIALGEEFVNK"
7117. 7587
/gene="argR"
/note="synonym: Z4596"
7117. 7587
/gene="argR"
/function="regulator; Amino acid biosynthesis: Arginine"
/note="Residues 1 to 156 of 156 are 100.00 pct identical
to residues 1 to 156 of 156 from Escherichia coli K-12
Strain MG1655: B3237"
/codon_start=1
/transl_table=11
/product="repressor of arg regulon; cer-mediated site
specific recombination"
/protein_id="AAG58365.1"
/db_xref="GI:12517858"
/translation="MRSSAKQBELVAFKALLKEEFSSQGEIVAAIQFGPDNLNQS
KVRMLTKFQAVTRNAKMEMVYCLPAELGVPVTPSSPLNKLVLDDYNDVAVVHTSP
GAAQLIARLDSLGKAEGLIGTIGADDTITFTTPANGFTVKDLYEAILLEFDQEL"
7901. 8215
/gene="yhcN"
Query Match 92.3%; Score 362.8; DB 1; Length 11904;
Best Local Similarity 95.5%; Pred. No. 4.8e-95;
Matches 364; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 TTCAACCTCAAAACGACAGTCGCGATATCAAAATAAAACAGCAGCAATAGAGCGCGGTGT 60

Db 11539 TTCAACCTCAAAACGACAGTCGCGATATCAAAATAAAACAGCAGCAATAGAGCGCGGTGT 11598
Qy 61 TGNACACCCCGGATGCCAGACAAAGTAGTCCTGTCAGATACCTGTCGCAACAGTACCCGCGCA 120
Db 11599 TGAACCAACCCCGGATGCCAGACAAAGTAGTCCTGTCAGATACCTGTCGCAACAGTACCCGCGCA 11658
Qy 121 CCAGCCAGAAATTCGCCAGTATATAAGCAATTCAAAAATATCGGTGGAAAGGACAGCC 180
Db 11659 CCAGCCAGAAATTCGCCAGTATATAAGCAATTCAAAAATATCGGTGGAAAGGACAGCC 11718
Qy 181 CAAACACCGAGTAACCGGAAACAGACTCATGTGACCTTGGTCTGTGTAAAGAGAGAGCAGG 240
Db 11719 CAAACACCGAGTAACCGGAAACAGACTCATGTGACCTTGGTCTGTGTAAAGAGAGAGCAGG 11778
Qy 241 CGTTATTATTTTCAGCATCTGTCGCCGACAGAGGCGCATGGAAGCCCGGCGGAGAGCAA 300
Db 11779 CGTTATTATTTTCAGCATCTGTCGCCGACAGAGGCGCATGGAAGCCCGGCGGAGAGCAA 11838
Qy 301 CATGCTGTAGATGATATTAATATATTAGCTAACTGTTATCGTGTATCTATATTAT 360
Db 11839 CATGCTGTAGATGATATTAATATATTAGCTAACTGTTATCGTGTATCTATATTAT 11898
Qy 361 GTGATC 366
Db 11899 GTGATC 11904
RESULT 8
AE008855 21252 bp DNA linear BCT 23-APR-2003
LOCUS Salmonella typhimurium LT2, section 159 of 220 of the complete
DEFINITION Genome.
ACCESSION AE008855 AE006468
VERSION AE008855.1 GI:16421920
KEYWORDS
SOURCE Salmonella typhimurium LT2
ORGANISM Salmonella typhimurium LT2
REFERENCE 1 (bases 1 to 21252)
AUTHORS McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W.,
Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F.,
Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A.,
Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W.,
Stonking, T., Nhan, M., Waterston, R. and Wilson, R.K.
Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2
JOURNAL Nature 413 (6858), 852-856 (2001)
MEDLINE 21534948
PUBMED 11677609
REFERENCE 2 (bases 1 to 21252)
AUTHORS The Salmonella typhimurium Genome Sequencing Project
CONSTRM Direct Submission
TITLE Submitted (29-MAR-2001) Genome Sequencing Center, Department of
JOURNAL Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT Supported by NIH grant 5U 01 AI43283
Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER; <http://www.tigr.org/soitlab/glimmer/glimmer.html> and
GeneMark; <http://opal.biology.gatech.edu/GeneMark/>
EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>,
and Pedro Romero and Peter Karp at EcoCyc;
<http://ecocyc.PangeaSystems.com/ecocyc/>
The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Vides and
ReguonDB;

http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

FEATURES

```
source
  Location/Qualifiers
    1..21252
      /organism="Salmonella typhimurium LT2"
      /mol_type="genomic DNA"
      /strain="LT2; SGSC 1412; ATCC 700720"
      /db_xref="ATCC:700720"
      /db_xref="taxon:99287"
      /note="LT2"
Gene
  207..482
    /gene="yhcn"
    /note="synonym: STM3361"
RBS
  207..212
    /gene="yhcn"
    /note="putative RBS for yhcN; RegulonDB:STMSIH003374"
  219..482
    /gene="yhcn"
    /note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 73% identity in aa 18 - 104"
    /codon_start=1
    /transl_table=11
    /product="putative outer membrane protein"
    /protein_id="AAL22230.1"
    /db_xref="GI:16421921"
    /translation="MKITVTATLSILSVLSFGFAAPPISAEQANREATIGSVSVA
    IGSSPMDNMAMLSKXADEQATAYHTEARSGSNWHATAELYK"
  579..852
    /gene="STM3362"
    /note="STM3362"
  579..584
    /gene="STM3362"
    /note="putative RBS for STM3362; RegulonDB:STMSIH003375"
  586..852
    /gene="STM3362"
    /note="similar to E. coli orf, hypothetical protein (AAC76270.1); Blastp hit to AAC76270.1 (104 aa), 53% identity in aa 18 - 104"
    /codon_start=1
    /transl_table=11
    /product="putative periplasmic protein"
    /protein_id="AAL22231.1"
    /db_xref="GI:16421922"
    /translation="MKTYIIASLGLATLLSFGANAAVHQNVAEQANLQPMGTISVS
    QIGSPMDMROEIVAKAKAGANSYRIIELKEDGNWHATAELYK"
  complement(912..1197)
    /gene="yhco"
    /note="synonym: STM3363"
    /complement(912..1184)
    /gene="yhco"
    /note="similar to E. coli orf, hypothetical protein (AAC76271.1); Blastp hit to AAC76271.1 (90 aa), 76% identity in aa 1 - 90"
    /codon_start=1
    /transl_table=11
    /product="putative cytoplasmic protein"
    /protein_id="AAL22232.1"
    /db_xref="GI:16421923"
    /translation="MNVYTFDFNDIKQSDFYREFTOTFGLASKEVSDLDLTLDVMS
    DILPLLEIFVHLDPKURRYGALILFDFAEELEGRLFRNVRH"
  complement(1192..1197)
    /gene="yhco"
    /note="putative RBS for yhcO; RegulonDB:STMSIH003376"
    /complement(1356..13323)
    /gene="yhcp"
    /note="synonym: STM3364"
    /complement(1356..13323)
```

```

/genes="yhcp"
/notes="similar to E. coli orf, hypothetical protein (AAC76272.1); Blastp hit to AAC76272.1 (655 aa), 92% identity in aa 1 - 655"
/codon_start=1
/transl_table=11
/product="putative inner membrane protein"
/protein_id="AAL22233.1"
/db_xref="GI:16421924"
/translation="MGIFSIANOHIRFAVKLACAIVLALFIGHFHOLETPRMAVLTA
IVAAGPAPAGGPEPSGAIRYRGMRLIIGTFIGCIAALIIISMIRAPLLMILVCCVQ
AGFTWISSIVRIENSYAGLSGYTALIIVITQTEPLTPQFALRCESEIVIGISCA
ILADLLFSPRSIRKQVDRELDLSLVAQYQIMQCIIRKGSSEVDNAGDLVRRTALE
GMRNLNMESSWRVRANRKAALNTLSLTITQSCETYLITQNTREPLITDTFRELFET
FVETQDVHRQKRMERIVWTGERETPTVTLXSWGATRYLLIKKGVISNTKISATE
EELIQGPVVKVESAEERHAMVNFRTLLSCILGTILMTWGTSGNAGMVAIVTS
LAMRLPNRMVCIDFIYGLAALGLLYPLVITPNTQOQMLLCLSLAVLGFPIGIE
VOKRILGSGALASTINIIVLDNPMTHFSQFLDSALGOIVGCMIAFIVILVRDKSK
DRTGRVLLNQFVSAAVSATNTVVRKRNRLPALYQQLFLLMKNKFGDLPKPRALITM
ITAQKRDRAPIPVNEDLSVFRHQRARTADHVISAGSDDKRRIRFGQLDELIDYOEK
LRIWEAPPQVTEPVKLTGLNKHQYQNALTDS"
/complement(3329..4270)
/gene="yhco"
/notes="synonym: STM3365"
/complement(3329..4261)
/gene="yhco"
/notes="similar to E. coli putative membrane protein (AAC76273.1); Blastp hit to AAC76273.1 (310 aa), 92% identity in aa 1 - 310"
/codon_start=1
/transl_table=11
/product="putative membrane located multidrug resistance protein"
/protein_id="AAL22234.1"
/db_xref="GI:16421925"
/translation="MKTLRKLSRTAITLVILAFIAIFRAWVYTESPWTRDARFS
ADVIAIPDVAGLITHVNDNOLVKKQDVLFITDOPRYOKALAEAEADVAVYQVLAQ
EKQEAQRNRNLGVOQMSREEDQANNVLOTVLHQAKAQTEDLAKLDERTVIRAP
ADGWNTLVNAYAGEFTIRGSTAVALKVKNISFYQVQAYMEETKLEGVPGVRAEITPLGS
NRVLKGVDSVAAAGTVINASTSDAKGATIDNSLWVRLAQRVFIIRLDEQQGNLWP
AGTTATVITGKQDRDASQDSFRLAHLRLRFBG"
/complement(4265..4270)
/gene="yhco"
/notes="putative RBS for yhcQ; RegulonDB:STMSIH003377"
/complement(4269..4481)
/gene="yhcr"
/notes="synonym: STM3366"
/complement(4269..4472)
/gene="yhcr"
/notes="similar to E. coli orf, hypothetical protein (AAC76274.1); Blastp hit to AAC76274.1 (90 aa), 98% identity in aa 24 - 90"
/codon_start=1
/transl_table=11
/product="putative inner membrane protein"
/protein_id="AAL22235.1"
/db_xref="GI:16421926"
/translation="MSLFPVIVVFLGSPFPPIFFELLLSLAFWLVRMLVPTGIYDFV
WHPALENTALYCLFLYLSRLFV"
/complement(4476..4481)
/gene="yhcr"
/notes="putative RBS for yhcR; RegulonDB:STMSIH003378"
  4654..5583
    /gene="yhcs"
    /note="synonym: STM3367"
    /complement(4654..5583)
    /gene="yhcs"
    /note="similar to E. coli putative transcriptional regulator LYSR-type (AAC76275.1); Blastp hit to AAC76275.1 (309 aa), 95% identity in aa 1 - 309"
    /codon_start=1
    /transl_table=11
    /product="putative LysR family transcriptional regulator"
```

```

/protein_id="AAL22236.1"
/db_xref="GI:16421927"
/translation="MERIKRMSVFAGKVEFGSFTAAAROLQMSYSSISQTVAKLEDL
QVKLNLRSTRISGILTEAGKIYQOCERMLHEVDVHOLYAFNNTPIGTRIGSSNM
AONVLAGUTAKLKEYPGLAVNLVTCIPAPLIADGLDVIRVQALQSSLSFRRLLGA
MMVYCAAKPYLAQVGEKPADJSSHSWLEYSVRPNPEFELIAPEGISIRLPIQGRF
VTNDPMTLVRMLTAGTGAYVPLMWVIDEINRGDLEILLPRYQSDRPVVALYTEKDK
LPLKVVVINALTDYFVDVAHLFGMHGRGKEK"
gene
  complement(5706..7151)
  /gene="tldb"
  /note="synonym: STW3368"

Query Match      73.7%; Score 289.8; DB 1; Length 21252;
Best Local Similarity 85.2%; Pred. No. 1.3e-73;
Matches 335; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

QY 1 TTCAACCTCAACAGCAGTCGCGATATCAATAAACAAGCAGCAATAGAGCGGGTGT 60
Db 4262 TTCAACCTCAACAGCAGCGCGATATTAATAAACAAGCAGCAATAGAGCGGGTAT 4321

QY 61 TGAACAACCGCGGATGCCACAGAAAGTCGTAGATACCTGTGGCAAGTAGTACCCGGCGCA 120
Db 4322 TGAACAAGCGCGGATGCCACAGAAATCATAGATCCCGTCGGCAAGCATCCGACGCA 4381

QY 121 CGAGCCAGAAATCCCGAGTGATAAAGCAATTCAAAATATCGGTGGAGGACGACGC 180
Db 4382 CCAGCCAAATAATGCCAGTGACAAAGCAATTCGAAAATATCGGTGGAGGAAAGAC 4441

QY 181 CAAACACCAAGTACAGCGGAAACAGACTCATGTTGACCTTGGTTTAAAGAGAGAGCAGG 240
Db 4442 CAAACACCAAGTACAGCGGAAACAGACTCATGTTGACCTTGGTTTAAAGAGAGAGCAGG 4501

QY 241 CGTTATATTTTTCAGCATCTGTCGGCGAGAGAGGATGGAAGCCGGCGGAGAGCAA 300
Db 4502 CGATAGATTGTTAAGC-TACATCGCGCAGAGAGGCAAGAGACGCCGCGAGAGCGCA 4560

QY 301 CATGTCGTAGATTGATATTTAATATATTAGGCTAACTGTTATGCTGTATCTATATAT 360
Db 4561 CATGTCGTATTCGATATTAATATATTAGGCTAACTGTTATGCTGTATCTATATAT 4620

QY 361 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
Db 4621 GTGATCTAAATCACTTTTAAGTCAGAGTGAATA 4653

RESULT 9
AL627278      258050 bp      DNA      linear      BCT 04-JUL-2003
LOCUS
DEFINITION
  Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18,
  complete chromosome, segment 14/20.
ACCESSION
  AL627278 AL513382
VERSION
  AL627278.1 GI:16504263
KEYWORDS
  Salmonella enterica subsp. enterica serovar Typhi
SOURCE
  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Salmonella.
  1 (bases 1 to 258050)
REFERENCE
  Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D.,
  Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G.,
  Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T.,
  Connor, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N.,
  Farrar, J., Felwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S.,
  Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P.,
  Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J.,
  Stevens, K., Whitehead, S. and Barrall, B.G.
  Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18
  Nature 413 (6858), 848-852 (2001)
JOURNAL
  MEDLINE 21534947
  PUBMED 11677608
REFERENCE
  2 (bases 1 to 258050)
  Parkhill, J.
  AUTHORS
  TITLE
  Salmonella enterica subsp. enterica serovar Typhi
  notes="PS00013 Prokaryotic membrane lipoprotein lipid
  attachment site"
  2434..3597
  /gene="STY3366"
  2434..3597
  /gene="STY3366"
  /notes="Fasta hit to YJFC_ECOLI (387 aa), 50% identity in
  392 aa overlap"
  misc_feature
  1835..1867
  /gene="STY3365"
  /notes="PS00013 Prokaryotic membrane lipoprotein lipid
  attachment site"
  2434..3597
  /gene="STY3366"
  2434..3597
  /gene="STY3366"
  /notes="Fasta hit to YJFC_ECOLI (387 aa), 50% identity in
  392 aa overlap"
  gene
  CDS
  
```

TITLE
JOURNAL

COMMENT

FEATURES
source

gene

CDS

gene

CDS

misc_feature

gene

CDS

Direct Submission
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
E-mail: parkhill@sanger.ac.uk

Notes:
Details of S. typhi sequencing at the Sanger Centre are available
on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S_typhi/).

Location/Qualifiers
1..258050

/organism="Salmonella enterica subsp. enterica serovar
Typhi"

/mol_type="genomic DNA"

/strain="CT18"

/db_xref="taxon:90370"

69..1544

/gene="tolC"

/note="synonym: STY3364"

69..1544

/gene="tolC"

/note="Similar to Salmonella enteritidis outer membrane
protein TolC precursor tolC SW:TOLC_SALEN (Q54001) (491
aa) fasta scores: E(): 0, 99.0% id in 491 aa, and to
Escherichia coli outer membrane protein tolC precursor
tolC or mtcb or muka or refi SW:TOLC_ECOLI (P02930) (495
aa) fasta scores: E(): 0, 89.7% id in 495 aa
Orthologue of E. coli tolC (TOLC_ECOLI); Fasta hit to
TOLC_ECOLI (495 aa), 90% identity in 495 aa overlap"

/codon_start=1

/transl_table=11

/product="outer membrane protein TolC precursor"

/protein_id="CAD07712.1"

/db_xref="GI:16504264"

/db_xref="GOA:Q8Z3N8"

/db_xref="TrEMBL:Q8Z3N8"

/translation="MOMKKLLPILGLSLGSGFTLSQAEMLMQVYQARLSNPERRKS
ADRDAAAFKINEARSPILPOLGLGADYTSNGYDANGINSNETSASLOLTOTLFDPM
SKRGWTLTLEKAGIQDVTYQDOTLIINTANAYFKVLNADIVLSYTAQKEALYRQ
LQCTOFNVGVAITDVQNAOYDTVLANEVTARNLNDNAVEIROVTGYPELA
SLNVEFKDZKXAVNALIKAEENSLLOARLSODIARBOICROAGHLPNLTA
STGIDTSYSGSKTNSAQIDDSNQCNGKIGLNFSLPQGGWNSQVQKAYNFVGAS
EQLESARHVSQVTVRSFNINASSINAYKQAVVQSSLDVAEAGTSVGTIRIVD
VLDTATLYDAKQQLANARYTYLINQNLKIYALGTINEQDLALANSTLIGKPIPTSPES
VAPETPEQDAADGYNAHSAPAVOPTAARANSNNNGPFRH"

1757..2428

/gene="STY3365"

1757..2428

/genes="STY3365"

/notes="Orthologue of E. coli ygiB (YGI_B_ECOLI); Fasta hit
to YGI_B_ECOLI (234 aa), 97% identity in 223 aa overlap.
Contains a possible N-terminal signal sequence."

/codon_start=1

/transl_table=11

/product="possible lipoprotein"

/protein_id="CAD07713.1"

/db_xref="GI:16504265"

/db_xref="TrEMBL:Q8ZG21"

/translation="MKRTKSIHHSFPRKSWARHLTPVALVAVFVFLAGCKSDSTV
SLYONADCSAANPGKSABCTAYNNALKEARTAPKYATREDVCAEFEGGCGQQA
QAGAPENQAQSSGSEFWMLMAGYMMGLMGAGGAGFACQPLPSSKNPASPAYGKY
TDAGKNYGAAPGRMTVFKTAMAPKPATTTVTRGGFSGVAKQSTMQRSAGTST
RSMGG"

1835..1867

/genes="STY3365"

/notes="PS00013 Prokaryotic membrane lipoprotein lipid
attachment site"

Orthologue of E. coli ygiC (YGIC_ECOLI); Pasta hit to YGIC_ECOLI (386 aa), 94% identity in 386 aa overlap

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07714.1"
/db_xref="GI:16504266"
/db_xref="TrEMBL:Q8FPF6"
/translation="MERSVITRDPDRKATEYGFNFTMTWGPYWCBDAYVYKLTAAQ
VEKLEDTAELQMCVKVVERVIADELMTKFIKPHTWGVRQSMOTQPSLSRLD
LAWDGIEPKLENNADPTSLYERAFQWLEQINAGNLPEGSDQPSLOEKLIE
RAERREQVQLHLHLCRDIREDRGTIQIQDCAEAETAEFLYIDDIQGLGKGG
FDLDQVLANFLKLYPWFEMFSTKLEDAWVLEPAWKSIIISNKALLPLLWEM
FDPHNLFAEEDPEHMDKYVVKPIREGANVSIENGKTTIESYEGYEGEEMI
VQQFPLPRFGDSYTLIGSLINDQAGIGIREDRALLITQLSREYPIHFVEG"
complement (3659..4489)
/gene="STY3367"
/complement (3659..4489)
/gene="STY3367"
/note="Orthologue of E. coli ygiD (YGID_ECOLI); Pasta hit
to YGID_ECOLI (271 aa), 87% identity in 266 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD07715.1"
/db_xref="GI:16504267"
/db_xref="GOA:Q8Z3N7"
/db_xref="TrEMBL:Q8Z3N7"
/translation="MVMILTCLSCDKNIMSLTCLPALFLGHGSPMVLDDNDYTRAWR
RLGEALPRQAIWVSAHWYTRGTGTAMERPTLHDFGPPQALYDWHYPAGSPAL
AORVELLAPVALDKAEMGPDHGSWGLIKWYNADIPMVOLSVDSTKPAAPHFEM
GRKALTRDEGVMVASGNVNLRTVRHGDNIPIYPAASFNDRFVKANLWQGPVEQ
HPLVNYLQEGGALNPTPEHFLPLLYLVIGAWDGRKEPITIPVDGIEMGSIISLMSVQVG"
4588..5361
/gene="STY3368"
4588..5361
/gene="STY3368"
/note="Orthologue of E. coli ygiE (YGI_ECOLI); Pasta hit
to YGI_ECOLI (257 aa), 93% identity in 257 aa overlap.
Contains multiple possible membrane spanning hydrophobic
domains"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAD07716.1"
/db_xref="GI:16504268"
/db_xref="GOA:Q8XGR4"
/db_xref="TrEMBL:Q8XGR4"
/translation="MSVPLILTLAAGATFAGFLVGLGKPSNRVLAFLSLGFAAGIM
LLISLMELPALDTEGMSPLVGYGMFTIIGLGYGLDELPHAPHDLPVOKROQPLP
GSIKETALLTIGLSHNPPEGIATFTVASSNLELGFGLALVALHNIPEGLAVGPFV
YATGSKRTAIFWAGISGMALIGGLVLAWLILGSLVSPVIVAAINAAVAGIWLVSVD
ELMPLAKSIDPNNSPSYVLCGMSLGLSLVILQITIGIG"
6004..7800
/gene="STY3370"
/note="synonym: asst"
6004..7800
/gene="STY3370"
/note="Similar to Klebsiella pneumoniae arylsulfate
sulfotransferase asst TR:P97036 (EMBL:U32616) (598 aa)
fasta scores: E(): 0, 86.8% id in 598 aa"
/codon_start=1
/transl_table=11
/product="probable arylsulfate sulfotransferase"
/protein_id="CAD07717.1"
/db_xref="GI:16504269"
/db_xref="GOA:Q8Z3N6"
/db_xref="TrEMBL:Q8Z3N6"
/translation="MFDQVRKTIILAGAVALTCLTAASTFAAGFOPAPQAGKLVAVVV
DPVGNAPTALVELDSHIIISDVKTVHKGKGVPTVTVGKESLETVDGPIPIGLYQ
KFNWTVVEYKENGKMKDDVYVQTSALVNHVMDNRSISDIQOKVIVKVAEGFEDRLY
LVNTHTFPQGAEPFHGKDKNAGILLDAGPAGGALPFDIATFTVVDITQGEYRWLLD

QDTFYNGHDMNINKRGYLMGIRETPRGTTAVQOQHWEYFDMQGOILADHKLPRGFID
ASHESIETVNGTVLLRVKRDYKEDGIHVHTIRDOIIEVDKSGRVVDVMDLTKILD
MRDALLGALDAGVACVUNVDLAHQOQAKSPDPTPYGDALGVAGRNWAFVNSIAYDAK
DDSIILSRHQGVKIGRQKWLAPSKGNWKLASKLLKPVDDHKGKPLTCDENGK
KDDTPDFTTQHTAMLSGKTLTVDFNDGRLGEPALFTMYKISRFVETKIDKGGT
VQVWEYKGRGDFYFSPITSVVEYQKDRMFGFGSINLFDVGKPTVGLKNEIDYK
TKYKVEIDVLSDGKPNQTHYRALLVHPQTQMKF"
7820..8491
/gene="dsbA"
/note="synonym: STY3371"
7820..8491
/gene="dsbA"
/translation="Similar to Enterobacter amnigenus disulfide
isomerase dsbA TR:Q9XDP1 (EMBL:AF012826) (222 aa) fasta
scores: E(): 0, 90.1% id in 222 aa"
/codon_start=1
/transl_table=11
/product="disulfide isomerase"
/protein_id="CAD07718.1"
/db_xref="GI:16504270"
/db_xref="GOA:Q8Z3N5"
/db_xref="TrEMBL:Q8Z3N5"
/translation="MSKWTITSLFKSVLTAALVTFFAASAFTEGTDYMLEKPIPA
DKTLIKVFSYACPCYKVDKAVTGPVSDKVADIVTPTPEHLETKGKQASVPAVL
LAKDKAAGISLFPDAKSOQKAKFAWYTAHDKERWSDKDPAAFIKTGLDAGMSQA
DFAALKDPAPVQETLEKWKAAAYDVAKIQGVPAVYVNGKYLIIYTKNIKSIDMAELVRE
LATKK"
7847..8488
misc_feature
7847..8488
Query Match 73.7%; Score 289.8; DB 1; Length 258050;
Best Local Similarity 85.2%; Pred. No. 1.5e-73;
Matches 335; Conservative 0; Mismatches 57; Indels 1; Gaps 1;
QY 1 TTCAACTCAAACGACAGTCGCGATATCAAAATAAAACAAGCAGCAATAGAGCGCGGTGT 60
DB 180504 TTCAACTCAAACGACAGCGCGATATTAATAAAACAAGCAGCAATACAGCGCGGTAT 180563
QY 61 TGAACACGCGGATCCGACAGCAAGTCGTAGTACCTGTGGCAGCAAGTACCGGCGCA 120
DB 180564 TGAACACGCGGATCCGACAGCAAGTCGTAGTACCTGTGGCAGCAAGTACCGGCGCA 180623
QY 121 CCAGCCAGAAATCGCCAGTCGATGATAAAGCAATTCGAAATAATTCGTTGGGAAAGCAGCC 180
DB 180624 CCAGCCAGAAATAGCCAGTCGATGATAAAGCAATTCGAAATAATTCGTTGGGAAAGCAG 180683
QY 181 CAACACACAGATAACCGGAAACAGACTCATGTTGACCTTGGTTGTAACAGAGAGCAGG 240
DB 180684 CAACACACAGATAACCGGAAACAGACTCATGTTGACCTTGGTTGTAACAGAGAGCAGG 180743
QY 241 CGTTATTATTTCAGCATCTGTCGCGCAGAGAGGSCATGGAAAGCGCGCGCAGAGCAA 300
DB 180744 CGATGATGTTTAAAGC-TACATCGCCGAGAGAGGSCAGAGAGCGCCAGCGAGAGCGCA 180802
QY 301 CATTCGTGTAGATGATTAATAATATATTAGCGTAACCTGTTATGCTGTTATCTATATTAT 360
DB 180803 CATAGCTCTTATTTCGAATAATAGTATTAACGTAATGTTGTTGCTGTTATCTATATTAT 180862
QY 361 GTGATCTAATCACTTTTAACTAGTCAGAGTGAATA 393
DB 180863 GTGATCTAATCACTTTTAACTAGTCAGAGTGAATA 180895
RESULT 10
AE016845
LOCUS
DEFINITION Salmonella enterica subsp. enterica serovar Typhi Ty2, section 12
of 16 of the complete genome.
ACCESSION AE016845
VERSION AE014613
KEYWORDS AE016845.1 GI:29139182
SOURCE Salmonella enterica subsp. enterica serovar Typhi Ty2
ORGANISM Salmonella enterica subsp. enterica serovar Typhi Ty2
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Salmonella.
 1 (bases 1 to 299991)
 Deng W., Liou S.R., Plunkett III, G., Mayhew, G.F., Rose, D.J.,
 Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R.
 Comparative Genomics of Salmonella enterica Serovar Typhi Strains
 Ty2 and CT18
 J. Bacteriol. 185 (7), 2330-2337 (2003)
 MEDLINE
 22531367
 PUBMED
 12644504
 REFERENCE
 2 (bases 1 to 299991)
 Deng W., Liou S.-R., Plunkett, G. III, Mayhew, G.F., Rose, D.J.,
 Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R.
 Direct Submission
 Submitted (25-SEP-2002) Laboratory of Genetics, University of
 Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
 Location/Qualifiers
 1. .299991
 /organism="Salmonella enterica subsp. enterica serovar
 Typhi Ty2"
 /mol_type="genomic DNA"
 /strain="Ty2"
 /serovar="Typhi"
 /sub_species="enterica"
 /db_xref="taxon:209261"
 complement (83..709)
 /gene="ftsJ"
 /locus_tag="t3214"
 complement (83..709)
 /genes="ftsJ"
 /locus_tag="t3214"
 /notes="corresponds to STV3475 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="cell division protein"
 /protein_id="AAO70750.1"
 /db_xref="GI:29139183"
 /translation="MTGKRSASVYVTOIGGKRIIACDLPMDIVGVDFLOQDFRD
 ELVKALLREVDGSKVQVYMSDAPNMGSTRAVDIPRAMYLVELALDNCRDVLAPGGS
 FVYKPGGSGFDEYREIKSLFKVKVRPDSRRSRREVIVATGRK"
 797..1129
 /locus_tag="t3215"
 797..1129
 /locus_tag="t3215"
 /notes="corresponds to STV3476 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAO70751.1"
 /db_xref="GI:29139184"
 /translation="WTRFOSQKQYTNWLSLKOKHLKGLAHPKPVVMLGNNGLTG
 GYLAIEEQALEHHELIKVKIASEDRETKLIIVDAIVRTGACNVQVIGKTLVLRPTK
 ERKISLPR"
 complement (1285..1761)
 /genes="greA"
 /locus_tag="t3216"
 complement (1285..1761)
 /gene="greA"
 /locus_tag="t3216"
 /notes="corresponds to STV3477 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="transcription elongation factor"
 /protein_id="AAO70752.1"
 /db_xref="GI:29139185"
 /translation="MQAIPMTLRGAELKEEEDFLKSVRRPEIIAIAEAREHGDLEKE
 NAEYHAARQQGFCEQKIDIAKLSNAQVIDVTKPNNRVRVIGATVTVLNLDDEE
 QTVRIYVGDEADPFQKNLSIVNSPIARGLGKQDDVWIKTPGGDVEYELKVEYL"
 2009..3442
 /locus_tag="t3217"
 2009..3442
 /locus_tag="t3217"

2009..3442
 /locus_tag="t3217"
 /notes="corresponds to STV3479 from Accession AL513382:
 Salmonella typhi CT18; D-alanyl-D-alanine
 carboxypeptidase"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein"
 /protein_id="AAO70753.1"
 /db_xref="GI:29139186"
 /translation="MFRSFIIGLTTSIAFSVQAANIDEYIKQLPAGANLALMVOKIG
 APAPDIYHSQWALPASTQKVTALAAALIQGPDPFRFTTILETKGVNDGILKGLDI
 ARFGDPTLRRODIRNMVATLKSGVTQIDGNVLDITSFASHDKAPGWNLDLQCF
 SAPAAAIYDRNCFSVLSAQKPNDLARIVASYVPTMFSQVRLTPEGSDADQYCE
 LDVVPGLNRYLTGCLPQADPLPAIQQGASVAGALLKQELKEAGITVYRGLLR
 QTCNVPEFTIVASKQAPLHDLKIMKKSDNMADIATVFRMIGHKRVNFVFGWRAGSD
 AVRIILQQAGIDIGNTIADGSGSRHNLIAPATMMQVLYIAQHDNELNFIIMPL
 AGYDGLQYRAGLHQAGVGVKSATGSLQGVYNLAGFTTASGORMAFVQYLSGVAV
 PPADQRNRIPLVRFESRLYKDIYQNN"
 complement (3573..4745)
 /locus_tag="t3218"
 complement (3573..4745)
 /locus_tag="t3218"
 /notes="corresponds to STV3480 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="probable GTP-binding protein"
 /protein_id="AAO70754.1"
 /db_xref="GI:29139187"
 /translation="MKFVDEASILVVAGDGGNGCVSFRREKYPKGGDGGDGGD
 VVMEADENLNTLIDYRFEKSFRAERGONGASRDCTGKRGKVTIKVPVTRVIDQGTG
 ETWGDMLTKGQRLVAKGWHGLGNTRFKSVNVRTPROKNTGTPGDKRDLLEMLLA
 DVGMLGMPNAGKSTFIRAVSAAPKVADYFTTLVPSLGVRMDSEKSPVADIPGLI
 EGAEAGAGLIRFKHLRCRVLHLLDIDIDIDGSDPYENARIIGELKYSQDLAAK
 PRWLNFNKIDMDKSEAEKAKIAEALGEGKRYIISASQGLKDLCDWNTFLIE
 NPITAEAEAKQKEVFEMDDYHQQLAEVEDADDWDDEDEDEGEVFIYKR"
 complement (4761..5726)
 /locus_tag="t3219"
 complement (4761..5726)
 /locus_tag="t3219"
 /notes="corresponds to STV3481 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="putative membrane protein"
 /protein_id="AAO70755.1"
 /db_xref="GI:29139188"
 /translation="MQQAGIGILLALTTCMCGALPIAMKQVLEWPESTIVYRFL
 MASIGILAVKRLPPIRPFKRWLVLAATCGLFGNFILFSSSLQYLSPTASQ
 VICQLSPGVWVASVFLIKERKQTVQIGALMLLSGVNFNTSLIEFTRLDYWG
 VIFGVGAAMVWSYGVAAQVLLRSLAQIILFTLTCTIALPLKPVIAQLSDWQ
 LACILFCGNTLVGYGALAEAVARWQAQVSAIITLPLTLLFSLLSWAMPDFPAR
 PMLNLLYGAFVUVVAGAMYSAIGHRIWGLRKHETVWSQPRSGE"
 complement (5856..6113)
 /genes="rpnA"
 /locus_tag="t3220"
 complement (5856..6113)
 /gene="rpnA"
 /locus_tag="t3220"
 /notes="corresponds to STV3482 from Accession AL513382:
 Salmonella typhi CT18"
 /codon_start=1
 /transl_table=11
 /product="50S ribosomal subunit protein L27"
 /protein_id="AAO70756.1"
 /db_xref="GI:29139189"
 /translation="MAHKAGGSTRNGRDSKRLGKRFGEAGVLAGSIIVRQRGTG
 FHAGTVGGRDHTLFAKADGKVKFVKPKRKYISIAVE"
 complement (6133..6444)
 /genes="rplU"
 /locus_tag="t3221"
 complement (6133..6444)

		/gene="rplJ"		/locus tag="t3221"		/note="corresponds to STY3483 from Accession AL513382: Salmonella typhi CT18"		/codon start=1		/transl_table=11		/product="50S ribosomal subunit protein L21"		/protein_id="AAO70757.1"		/db_xref="GI:29139190"		/translations="MTAVFSGGQKQHRVSEGGQVRLEKLDIATGETIEFAEVLMIANG EYVKIGVPFVDGVIKAEVVAHGKEKVIKVPFRKHYKQKQGHQWFTDVKITGIS A"									
gene		6703..7674		/gene="ispB"		/locus tag="t3222"		6703..7674		/gene="ispB"		/locus tag="t3222"		/note="corresponds to STY3484 from Accession AL513382: Salmonella typhi Crl8"		/codon start=1		/transl_table=11		/product="octaprenyl-diphosphate synthase"		/protein_id="AAO70758.1"		/db_xref="GI:29139191"		/translations="MLEKINELTAQDMAGVNAITLEQLNSDVQLINOLGYVILSGGG KRIIPAVLAAKAVCYQGNHYTIAALIEFTHTATLLHDDVDSMDMRGKATANRA FQNAASVLGVDFIYTFRAFQMTSLGSLKVLVEVNSEAVVIAEGVLTQLMNVNDPDI EENRVISYKGTFFATGAAACQCSGILAGCTPEQKGLQDYGRYLGTAFQILDLLDYSA DGHGLGNVDDINEKQPTLPLHANMRHGTPEQSAMIRTAIEQNGRHLLLEVPLEAMT CCSLEWTRQABEEADKAISALQILPDPWREALTGLAHIAVQRDR"	
gene		7907..8194		/gene="nlp"		/locus tag="t3223"		7907..8194		/gene="nlp"		/locus tag="t3223"		/note="corresponds to STY3485 from Accession AL513382: Salmonella typhi Crl8"		/codon start=1		/transl_table=11		/product="Ner-like regulatory protein"		/protein_id="AAO70759.1"		/db_xref="GI:29139192"		/translation="MESKLIIDWHPADIIAGLRKKGTSMVAESRRNLSSSTILANALTR 73.7%; Score 289.8; DB 1; Length 299591; Best Local Similarity 85.2%; Pred. No. 1.5e-73; Matches 335; Conservative 0; Mismatches 57; Indels 1; Gaps 1;	
QY	1	TTCAACTCAACGAAACAGTCGCGATATCAATAAACAAGCAGCAATAGCGCGGTGT 60																									
	Db	65335		TTCAACTCAACGAAACAGCGCGATATTAATAAACAAGCAACAATACACGCGGTAT 65394																							
QY	61	TGAACACGCGGATCGCAGACAAAGTCCTAGTACCTGTTGGCACAAGTACCCGGCGCA 120																									
Db	65395	TGAACAGCGGATGCCAGACAAATCATAGATGCCCTCGGCAACAGCATCCGACGCA 65454																									
QY	121	CCAGCCAGAAATTCGCCAGTGAATAAAGCAATTCAAAAAATATCGGTGGGAAGGACAGCC 180																									
Db	65455	CCAGCCAAAAATAGCCAGTGACAAAGCAATTCGAAAAATATCGTGGGAGGAAAGAC 65514																									
QY	181	CAAACACCAAGTAACGGGAACAGACTCATGTTGACCTTGGTTGTAAGAGAGACAGG 240																									
Db	65515	CAAACACCAAGTAACGGGAACAGACTCATGTTGACCTTGGTTGTAAGAGAGACAGG 65574																									

Db	11579	TGACACACGGCGGATGCGACACAAAGTCGTAGTACTCTGTGGCAACAAGTACCCGCGCA	11638	CDS	complement(1285..2730) /gene="tldb" /locus_tag="SF3283" /note="Residues 1 to 481 of 481 are 100 pct identical to residues 1 to 481 of a 481 aa protein from Escherichia coli K12 ref: NP_417711.1" /codon_start=1 /transl_table=1 /product="suppresses inhibitory activity of CsrA" /protein_id="AAN44747.1" /db_xref="GI:24053721" /translation="MSLNLSVEQLLAANGKHQDLFAILQQLAERRLDYGDLYPQSSY HESWLEDRILKDSYNIQDQGVRAISGKTFAYADQISLLALEQAQAARTIVR SDGKQVTLGAHSHPLTSDPLQMSREEKLDILRRVDKVAEADKRVQEVNTASLS GYELIIVAAATDGTLAADVRPLVLSVLEEDKGERGASGGGFGFVEFLADLD GEDVADAKAEAVMALNLSAVAAPAGTVPVLGACWPGVLLHEAVGHLEGDFFNER GTSVFGQVGLVASLCTVVDGTVWDRGSAIDDEGTPGQVNLLENGILKGYNQ DKLNARLMGTPPGNGRRSYAHLPRMPTNTYMLPKSTPQEIIEESVGIYAPNFG GGQVDITSGKRFVSTSEAYLIENGKVTKPVGATLIGSGIETWQIQSMVGNLKLONG VVGCGEGQLPVGVGQPTLLKVDNLTVGGTA"
Qy	121	CGAGCCAGAAAATCGCCAGTGATATAAAGCAATTCAAAATAATCGGTGGGAAGCAGCC	180	gene	/locus_tag="SF3284" /gene="yhdRp" /locus_tag="SF3284" /note="Residues 1 to 1266 of 1266 are 99 pct identical to residues 1 to 1266 of a 1266 aa protein from Escherichia coli K12 ref: NP_417712.1" /codon_start=1 /transl_table=1 /product="Uncharacterized membrane protein" /protein_id="AAN44748.1" /db_xref="GI:24053722"
Db	11639	CGAGCCAGAAAATCGCCAGTGATATAAAGCAATTCAAAATAATCGGTGGGAAGCAGCC	11698	CDS	/translation="MRRPLGILLTGAALVIAALLVSGLRILALPHNAWRPEILNKI ESATGMPVEASQLSAQNFQPTLEAHDIRAEUKDGGERSVKRVTLALDVQSLHMR WORDELTFWOLRERTNTPITSGGSDSLASHISDLRFQFCHFDLSDSVSLTSPG ORALELAPQLTWLNDPRRRAEGLVSLISITGQGVQWVRMDLRDEGLSNGRWLQK ADIDILKPLGKWMQDNIALETAQFSGEWMITDKGVDVGGVNLQGGASWGEKQK HTLVNLTVAHITRENPGQFSIPDTRITMDGKWPSPGALTAWIPEQDVGGKDKRS DELIRASNLLEAGLEGIRPLAAKSPALGDVWRSTQPSGKINTALDIPLOAADKTR FQASWDLAWKQWKLIPGAHFSGTSGSVENGLLTASMKQAMPVETVFRAPLEIAD QATISGLNNDKGQFLDGRNIDVAKAFAHAGGRFYLOPANDPEWLGILAGTDDGS QAWRYPNLNGKDLVDYLSGATOGGEADNATLVGNGNQLPEYKNEGQFVLPRLR NAKFAEQPWPAITNLDIELDFINDGLWKTGQVNLGVASNLTAIVTIDYSKEKLLI DADIKGPRGAVGYDETLEKSLQGLQELQDGDVNRALHLDIPLNGLVTAKEGV TLRNLSLFTKPLDLSILNLSGKFSFNGDLQSEPTASFNQPLNVDFSTKEGAKAYQ VAYNLNGWQPAKTGYLPEAVNEALSQSVAMDKGVGIDILPYHAGATYVELNGDLNVY SSHLPSPLAKPAGEPLAVNVKVDNLSNLSPELTQAGADNHSRWLLGQKLTLDRAIW AADSKTLPPLPEOSGVELNMPMNGAEWLALFKGAESVGGAAAPQHITLPTMLS LGNQOWNLSIVSQPTANGTAVEAGREINATLANNENAPLANIKYLYNPSVAKTR GDSFRSPPTTERINFRWPDAQIRCTECWQKFGKGRIDSDITISGOTLITLNGLI DTGSLTADGGEWNNPNERTSLKGRKRGKIDAAAEEFFGVTTPIRQSFNVYDLH WRKAPWQDEATLNGIHTOLGKEITEINTHAGQLLRLSLVDALMKRLRFDRTF GEGFYDSITAWIKDGMHTDDTLVDGLEADIAKMGSVNLVRDLNMEAVVAPIS ATVGAFAAFVNPVGAFAAFAASKVLGPLWSKSVILRYHISGLDDEPQINEVLRQPRK EKAQ"
Qy	181	CAAAACACCAGATAACGGGAACAGACTCATGTGACCTTGTTGT	226	gene	complement(6754..8241) /gene="cafa" /locus_tag="SF3285" /complement(6754..8241) /gene="cafa" /locus_tag="SF3285" /note="Residues 1 to 495 of 495 are 99 pct identical to residues 1 to 495 of a 495 aa protein from Escherichia coli O157:H7 EDL933 ref: NP_289814.1" /codon_start=1 /transl_table=1 /product="bundles of cytoplasmic filaments" /protein_id="AAN44749.1" /db_xref="GI:24053723"
Db	11699	CAAAACACCAGATAACGGGAACAGACTCATGTGACCTTGTTGT	11744	CDS	/translation="MRKGINNTAELLVNTPSETRVAYIDGILQEIHERARGIV GNLYKGRVSRVLPQAAFLIGLDRKAAPFLHASDIMPTECVAGSEQKQFTVRDISEL
RESULT 12	AE015339	Shigella flexneri 2a str. 301	linear	BCT 18-OCT-2002	
LOCUS	AE015339	Shigella flexneri 2a str. 301	DNA	section 302 of 412 of the complete genome.	
DEFINITION	AE015339	Shigella flexneri 2a str. 301			
ACCESSION	AE015339	Shigella flexneri 2a str. 301			
VERSION	AE015339.1	Shigella flexneri 2a str. 301			
KEYWORDS	GI:24053719	Shigella flexneri 2a str. 301			
SOURCE		Shigella flexneri 2a str. 301			
ORGANISM		Shigella flexneri 2a str. 301			
REFERENCE		Bacteria, Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Shigella.			
AUTHORS		1 (bases 1 to 11734) Jin, Q., Yuan, Z. H., Xu, J. G., Wang, Y., Shen, Y., Lu, W. C., Wang, J. H., Liu, H., Yang, J., Yang, F., Qu, D., Zhang, X. B., Zhang, J. Y., Yang, G. W., Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P., Kan, B. T., Chen, S. X., Yao, Z. J., He, B. K., Chen, R. S., Ma, D. L., Qiang, B. Q., Wen, Y. M., Hou, Y. D. and Yu, J.			
TITLE		Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157			
JOURNAL		Nucleic Acids Res. 30 (20), 4432-4441 (2002)			
PUBMED		12384590			
REFERENCE		2 (bases 1 to 11734) Jin, Q., Shen, Y., Wang, J. H., Liu, H., Yang, J., Yang, F., Zhang, X. B., Zhang, J. Y., Yang, G. W., Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P., Chen, S. X., Yao, Z. J., Wang, Y., Lu, W. C., Qiang, B. Q., Wen, Y. M. and Hou, Y. D.			
AUTHORS		Direct Submission			
TITLE		Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry of Public Health, 100 Yingxin Jie, Xuanwu Qu, Beijing 100052, P.R. China			
JOURNAL		Location/Qualifiers			
PUBMED		1..11734			
REFERENCE		/organism="Shigella flexneri 2a str. 301" /mol_type="genomic DNA" /strain="301" /serotype="2a" /db_xref="taxon:198214" 228..1157 /gene="yhcs" /locus_tag="SF3282" 228..1157 /gene="yhcs" /locus_tag="SF3282" /note="Residues 1 to 309 of 309 are 99 pct identical to residues 1 to 309 of a 309 aa protein from Escherichia coli O157:H7 ref: NP_312143.1" /codon_start=1 /transl_table=1 /product="putative transcriptional regulator LYSR-type" /protein_id="AAN44746.1" /db_xref="GI:24053720"			
gene		translation="MERLKRMSVFAKVVEFGSFTAAARQLQMSVSSISQTVSKLEDEL QVNLNRSTRIGLTAAGRIYQGRMLHEVDVHQLVAFNFTPIGTLRIGCSSFM AQNVLAGTAKMLKEYPELSANLVGTIPAPDLIADGLDVTIRVICALODSSILFSRLGA MEMVCAKSLYTOYGIPEKPADLSSHWLEYSVRPNPELIIAGPIEGISLRLFPQGRF VTNDPMTLYRLTLAGATAYVPLMWVINEINRGEILLPRYQSDPRPVLYTEKDK LPLKQVINSILDTIFVEVGKLFQEMHGRGKEK"			
CDS		complement(1285..2730) /gene="tldb" /locus_tag="SF3283"			
gene					

```

VRQGLDMVQVVKPLGTGKARLTDTILPSRYLVFMGASHGVSVORIESESEREL
KKVAAYDEOQGFIRTAAGVGVGAELASDAAYLKRVTXVMERKKRPOTRYQLYGE
LALAQRVLDFADALDIRVDSRLTYEALLSEYIPWMTSKLHVTPGPIFDLF
DVENIQALREKVELKSGVLIIDQTEAMTVDINTGAFVGRNLDITPNTWIEAT
QAIARQLRNJGGIIDIIDMNNEDHRRVRLSHLEQALSCKDRVKSIVNGFSALGV
ENRKRTRRESIBHVJCNCPCHGTGRTVETVCYEMREIRVIRVHAIDSDRFLVTA
SPVAEALKGESHSALAEVIFVQKQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV
complement (8213..8806)
/gene="yhde"
/locus_tag="SF3286"
complement (8213..8806)
/gene="yhde"
/locus_tag="SF3286"
/notes="Residues 1 to 197 of 197 are 97 pct identical to
residues 21 to 217 of a 217 aa protein from Escherichia
coli gb: AAS8051.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AAN44750.1"
/db_xref="GI:24053724"
/translation="MTSLYLAGSPRRQELLAQLGVTFEIVTGIEEQRPQBSAQY
VRLAREKAQAGVAGTACLVPLVADIIIVLNGEVLKPRDAEHAQMLRLSGTHQ
VMTAVLADQSHILCLVVDVTFRTLTDEDIAGVVASGEPLDKAGAYIQGLGCFV
RKINGSYHAVGLPLVEFYELLSNPNALREKDKDG"
complement (8815..9303)
/gene="mreD"
/locus_tag="SF3287"
complement (8815..9303)
/gene="mreD"
/locus_tag="SF3287"
/notes="Residues 1 to 162 of 162 are 98 pct identical to
residues 1 to 162 of a 162 aa protein from Escherichia
coli O157:H7 EDL933 ref: NP_289816.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AAN44751.1"
/db_xref="GI:24053725"
/translation="WASYRSQGRVWISFLIALLQLQMPEDNLIVFRPNWLLILL
YMLALPHRVNVTGTFWMAIILDSGMLGVRLVMSIIAYLVALKYOLFRNLALWQ
QALVVMLLSLVDIIVFAEFLVNVSEFVSWVNGVLPWIPFLMRKVRQQA
VQ"
complement (9303..10406)
/gene="mreC"
/locus_tag="SF3288"
complement (9303..10406)
/gene="mreC"
/locus_tag="SF3288"
/notes="Residues 1 to 367 of 367 are 99 pct identical to
residues 1 to 367 of a 367 aa protein from Escherichia
coli K12 ref: NP_417716.1"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein"
/protein_id="AAN44752.1"
/db_xref="GI:24053726"
/translation="MKPIFSRGPISLOIRLILAVLALGIIIDSLRGLTFSQIRTYMDT
AVSPFVSNAPRELLDGI SOTLASDQLELENRALRQELLNKSELMLGQYKQENA
RLRELLGSPRDEQKQVQVISTVNDPSQDVQVQVSGVSGQVPSKGVVQ
VVAVALTRVLLICATHALPIQVLRNDIRVAGNCTDQLQELHLPANTDIRVGD
VLVTGLGRFPFEGYVAVSVYKLTQRAYTVIQARPTAGLQRLIRLLMLGARDNG
ENPMTEEVHRVANERLMQVQLSPDMGPKLPATGATGATPQOPATGNAATA
PVAPTQPAANRSPQATPQSQAPPARAPGG"
complement (10472..11590)
/gene="mreB"
/locus_tag="SF3289"
complement (10472..11590)
/gene="mreB"
/locus_tag="SF3289"
/notes="57.4%; Score 225.4; DB 1; Length 11734;
Best Local Similarity 99.6%; Pred. No. 9e-55;
Matches 226; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

167 TGGGAGGAGCGCCCAACACGACGATAACGGGAACAGACTCATGTTGACCTTGTGTTGT 226

```

```

1 TGGGAGGAGCGCCCAACACGACGATAACGGGAACAGACTCATGTTGACCTTGTGTTGT 60
227 AAAGAGAGAGCAGCGCCTTATTATTTTTCAGCATCTGTGCGCGAGAGAGGGCATGGAAG 286
61 ATAGAGAGAGCAGCGCCTTATTATTTTTCAGCATCTGTGCGCGAGAGAGGGCATGGAAG 120
287 CCGGGGAGAGCAGCAATCTGTAGATTGATATTTAATATATATAGCTTAACCTGTTATGCT 346
121 CCGGGGAGAGCAACATCTGTAGATTGATATTTAATATATATAGCTTAACCTGTTATGCT 180
347 GTTATCTATATTATGTGATCTAAATCACTTTTAAGTCAGAGTGAATA 393
181 GTTATCTATATTATGTGATCTAAATCACTTTTAAGTCAGAGTGAATA 227

RESULT 13
AR389738/c
LOCUS AR389738 258 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6467 from patent US 6610836.
ACCESSION AR389738
VERSION AR389738.1 GI:40099919
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 258)
AUTHORS Breton, G.L. and Osborne, M.
TITLE Nucleic acid amino acid sequences relating to Klebsiella pneumoniae
for diagnostics and therapeutics
JOURNAL Patent: US 6610836-A 6467 26-AUG-2003;
FEATURES Location/Qualifiers
source 1..258
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 37.3%; Score 146.6; DB 6; Length 258;
Best Local Similarity 78.3%; Pred. No. 8.5e-32;
Matches 176; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 8 TCACACGACAGTCGCGATATCAATAAACAACAGCAGCAGCAGCAGCGGTGTTGAACA 67
DB 258 TCACACGACAGTCGCGATATCAATAAACAACAGCAGCAGCAGCAGCGGTGTTGAACA 199
QY 68 CGCGGATGCCAGACAAAGTCGTAGATACCTGTGTCACAAAGTACCCGCGCACCAGCCA 127
DB 198 TCGTGGTGGCCAGACAAATCGTAAATGCCGTCGGAACCCAGCAGTCGCTGACACGCA 139
QY 128 GAAATCGCAGTGATAAAGCAATTCAAAAATATTCGTGGGAGAGCAGCAGCAGCAACAC 187
DB 138 AAAGATAGCCAGCGCACAAAGCAACTCGAAGAGATCGCGGAGAAAGATTAAGCGGAAAT 79
QY 188 CACGATACCGGAAACAGACTCATGTTGACCTTGTGTTGTAAGAG 232
DB 78 CACGATACCGGAAACAGACTCATGATGACCTTGAATAAGACG 34

RESULT 14
AR389738/c
LOCUS AR389738 593 bp DNA linear BCT 01-MAR-2002
DEFINITION Serratia marcescens yhcR gene for hypothetical protein, complete
cds.
ACCESSION AB080601
VERSION AB080601.1 GI:19032360
KEYWORDS
SOURCE Serratia marcescens
ORGANISM Serratia marcescens
REFERENCE 1
AUTHORS Baba, M., Midorikawa, Y., Nakagawa, Y., Fujita, M. and Matsuyama, T.
TITLE Serratia marcescens and Escherichia coli genes controlling

```


REFERENCE	JOURNAL	AUTHORS	TITLE	JOURNAL
2	(bases 1 to 593)	Nakagawa,Y., Baba,M. and Matsuyama,T.	Direct Submission	
Submitted	(27-FEB-2002)	Yoji Nakagawa, Niigata University, Faculty of Agriculture, 8050,Ikarashi 2-no-cho, Niigata City, Niigata 950-2181, Japan (E-mail:nakagawaagr.niigata-u.ac.jp, Tel:81-25-262-6360(ex.6360), Fax:81-25-262-6360)	Location/Qualifiers	
1.	593		/organism="Serratia marcescens"	
/mol_type="genomic DNA"			/db_xref="taxon:615"	
301..504			/gene="yhcr"	
301..504			/codon_start=1	
/transl_table=1			/product="hypothetical protein"	
/protein_id="BAB95653.1"			/db_xref="GI:19032361"	
/translation="MSLFPVMVIFGLSPFPVFELLVPLALFLLRLLQPTGIYDFV"				
ORIGIN				
Query Match	32.0%	Score 125.8;	DB 1;	Length 593;
Best Local Similarity	75.1%	Pred. No.1.1e-25;		
Matches 15;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;
QY	3	CAACCTCAAAAGCAACAGTCGCGATATCAAAATAAAACAGCAGCAATAGAGCGGTGTG	62	
Db	509	CGACCTCAAAACAAAAGACATGAATCAGATPAAACAGCAGCAATACAGCGCAGTGTTA	450	
QY	63	AACACCGCGGATCCGACAGAACTCGTAGTACTGTGTGCACAAGTACCCGGCGCAC	122	
Db	449	AACAGCGCGGATCCGACAGAAATCGTAATCCCGCTCAGCAGCGCGCGCAAG	390	
QY	123	AGCCAGAAAATCGCCAGTGATAAAGCAATCAAAAATAATTCGTTGGGAAGGACAGCCCA	182	
Db	389	AGAAAACAGCGCGCAGCGCACAGCTCGAAAACCCGCGGACACAGTCCG	330	
QY	183	AACACAGCATACCGGAAACAGACTCAT	211	
Db	329	AAGATGACCATACCGGAAGCAACTCAT	301	
RESULT 15				
AE013617				
LOCUS	AE013617	10553 bp	DNA	linear BCT 26-JUL-2002
DEFINITION	Yersinia pestis KIM section 17 of 415 of the complete genome.			
ACCESSION	AE013617	AE000952		
VERSION	AE013617.1	GI:21956840		
KEYWORDS				
SOURCE	Yersinia pestis KIM			
ORGANISM	Yersinia pestis KIM			
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
Enterobacteriaceae; Yersinia.				
1 (bases 1 to 10553)				
Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,				
Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,				
Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,				
Straley,S.C., McDonough,K.A., Nilles,M.L., Matson,J.S.,				
Blattner,F.R. and Perry,R.D.				
Genome Sequence of Yersinia pestis KIM				
J. Bacteriol. 184 (16), 4601-4611 (2002)				
12142430				
2 (bases 1 to 10553)				
Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,				
Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,				
Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,				
FEATURES				
source				

```

/codon_start=1
/transl_table=11
/product="putative ribonuclease"
/protein_id="AA083768.1"
/db_xref="GI:21956844"
/translation="MNKELMALIGAVLLIAGAQGIDGTVTRDVGSSAELERPALSA
SAATESLTHQKVKYKLTHTHRLPDFVLTKEQBEQGWDAKGNLCVLPKATGGDR
FENRRKQPLDAGRWRRADRYRCGRGRSRLYSNDGLIYLTDHYKHPIME"
/complement(3264..3602)
/gene="y0175"
/complement(3264..3602)
/gene="y0175"
/function="unknown"
/notes="residues 1 to 112 of 112 are 59.82 pct identical to
residues 374 to 485 of 486 from GenPept : >gb|AA044216.1|
(AE009270) succinate semialdehyde dehydrogenase
[Agrobacterium tumefaciens str. C58 (U. Washington)]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AA083769.1"
/db_xref="GI:21956845"
/translation="MIADNEQMKAASEETGPIAACRFRFTEAEVIOQANNTPEGLA
AYFTYQNLQVRVSEALPESGMIGVNECSVSTELAPFGGVKESGLGREGSVLGLDFEM
EVKTLHLGNL"
/complement(3599..4732)
/gene="gabD"
/notes="synonym: y0176"
/complement(3599..4732)
/gene="gabD"
/function="enzyme; central intermediary metabolism: Pool,
multipurpose conversions"
/notes="NADP-dependent activity; residues 23 to 366 of 377
are 56.06 pct identical to residues 15 to 360 of 482 from
E. coli K12 : B2661"
/codon_start=1
/transl_table=11
/product="succinate-semialdehyde dehydrogenase"
/protein_id="AA083770.1"
/db_xref="GI:21956846"
/translation="WPAONPLOHQDHGIYHIGVGVGKWSQAQEQFDVYVPATGEL
VARVAKSGQFTEAIKAASEAPFAWRKTPAKORAEILQRWVLLIMEHQSLAELMVS
EOKPKDELVEVAATASTQFSEQAKRANGELIIPSAKEGARILATREPIGVVAIT
PNWFLPMUTRKLGPALAGCTGLIKPANNPTLSAFALLALAEQAGVPAGVINGVGD
THAISDAIMASSDVRKISFTSTNGVTKLMRNAATMKRI SNELGNNAPYIVFDADL
EAAVAGAMACKVRNAGQVCVNRVFIQDAIYDEFVNLRAAEVKKLVGNGMDKDVNM
GPLINLAGLEKVEEHVDALKEGGRLLAGGIAMRWAVISSOR"
/complement(4999..6954)
/gene="y0177"
/complement(4999..6954)
/gene="y0177"
/notes="residues 9 to 650 of 651 are 64.64 pct identical to
residues 11 to 652 of 655 from E. coli K12 : B3240;
residues 9 to 650 of 651 are 64.79 pct identical to
residues 11 to 652 of 655 from GenPept :
>gb|AA058368.1|AE005551.1(AE005551) orf, hypothetical
protein [Escherichia coli O157:H7 EDL933]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA083771.1"
/db_xref="GI:21956847"
/translation="MTHPSFIRLRPAKLSFAIYVAALFLGFHLQLETPRWSVLTAAIV
SAGPAFAAGPEPSGAIRKRWMLRIGTFIGGIGLVILVITIRAPVLTMLCCLWAG
ICTWISLVREVSAYFAGLAGYATLIIIVTTGETPLLPQFAVERCSEIVIGVCANM
ADLLFSPRIKQDIDRLVDKLVQDQYLLQCIQPAEKSEIDRAWNELVKNITSLNGM
RSLVLMWESRQRCNRRLQVLTHTESLALITQACETVLYNSNHPVISAELKMLSEPA
OTPAETHQOKMLRQFIASHGSEAI PHTI SSWVGAATRYLLLSKGIQTNINSQVEED
ILAGDAPVPIAEGHGHAMNGLRTGIATAGLGFMLWTGWTSGAGCMVIAVTSLSA
MRTPNRRALDFLVGVIIALPIGALYFMFIIPSTQCSMLLILCISGLVLAFTIGIEVQ
KRRIGSLGTASTINIVLSNPMIFNVRQFLDSALGQIVGCVFVSLVILLITPDNAKDR
TGTLLNRRFVSAYSAITNKTGRGENHLPALYQLNQLMMFPADIDIKRIALLIIL
AHQRLNRTEIPVNAELSAFHKQIRSAERVIIVNNDQKRYFYFARLLQELDQYQKLV
```

```

DYQAAADVIRPVKRLTEMLRKYSALI"
/complement(6956..7891)
/gene="y0178"
/complement(6956..7891)
/gene="y0178"
/function="putative membrane"
/notes="residues 1 to 311 of 311 are 72.66 pct identical to
residues 1 to 310 of 310 from E. coli K12 : B3241;
residues 1 to 311 of 311 are 100.00 pct identical to
residues 1 to 311 of 311 from GenPept : >emb|CAC93154.1|
(AJ414158) putative HlyD family secretion protein
[yersinia pestis]"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="AA083772.1"
/db_xref="GI:21956848"
/translation="MSTFSLKIIRVQITVVLVAVIAIFNVWAFVETSPWTRDAKFT
ADVVAIPDVSGLLTEVPKONQLVKQGLIFVIDQPRYQQAALAEADVAAYVOTLAA
EKQRESSRHLGIQALQOEIDQASNVLTQVHQAKAIAVRDLARLDLERTTVRAP
ABGWVTLNVHAGEFINRGATVALVKDITFYILAYLEETKLEGVKPGFRAETIPGS
NRLHGTVDISIGAVTINSSSADSKGLATIDNNLEWRLAQRPVVKIHLDSDDQOYLY
PAGTTATVITIGDNRDPHCASPMTKLMHLREFG"
/complement(7899..8102)
/gene="y0179"
/complement(7899..8102)
/gene="y0179"
/notes="residues 1 to 66 of 67 are 81.81 pct identical to
residues 24 to 89 of 90 from E. coli K12 : B3242"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

```

```

Query Match      29.9%; Score 117.4; DB 1; Length 10553;
Best Local Similarity 72.9%; Pred. No. 4e-23;
Matches 151; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy      5  ACTCAACGACAGTCGCGATATCAATAAACAAGCAGCAATAGAGCGCGGTGTTAA 64
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      7896  ACCTCAGGAAAAAGACGGGATGTTAAATAAACAGCAACATACAGCGCGGTATTA 7955

Qy      65  CAACGCGCGATGCCAGCAAAAGTCGTAGATACCTGTTGGCACAACTACCCGCGCACAG 124
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      7956  AAGGCGCGGTGCCAAACAACTCGTAATACCGTGGTTCGAGGATGCGCGCACAC 8015

Qy      125  CCGAAATCGCAGTGATAAAGCAATTCAAAATAATCGGTGGGAAGGACAGCCCAA 184
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      8016  AAAAAACAAAGCCCAATGAATATTAGCAGCTCAAGAAATATCGGTGGAAAGAGAGGCCAAA 8075

Qy      185  CACCAGCATACCGGAAACAGACTCAT 211
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      8076  AATCACCATCACCAGGACAACTCAT 8102
```

Search completed: December 19, 2004, 01:09:55
Job time : 1930.01 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 20:46:30 ; Search time 257.841 Seconds
(without alignments)
8001.134 Million cell updates/sec

Title: US-10-759-889-3

Perfect score: 393

Sequence: 1 ttcaacctcaaacgaacagt.....cttttaagtcagagtgaata 393

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	146.6	37.3	258	11 ABD00692	Abd00692 Klebsiell
2	100	25.4	100	8 ACD69190	Acg69190 E. coli K
3	100	25.4	100	8 ACD69191	Acg69191 E. coli K
4	100	25.4	100	8 ACD69192	Acg69192 E. coli K
5	38.4	9.8	1181	2 AAV12742	Aav12742 RBE1 tran
6	37.2	9.5	1035	11 ACH95383	Ach95383 Klebsiell
7	37	9.4	35000	6 ABT12817	Abt12817 Human REC
8	36.8	9.4	13503	6 AB134071	Ab134071 Human imm
9	36.6	9.3	397	4 AAI16233	Aai16233 Probe #61
10	36.6	9.3	397	4 ABA58948	Aba58948 Human foe
11	36.6	9.3	397	4 AAI38665	Aai38665 Human foe
12	36.6	9.3	397	4 ABA27808	Aba27808 Probe #73
13	36.6	9.3	397	4 AAK32852	Aak32852 Human bon
14	36.6	9.3	397	4 AAK07110	Aak07110 Human bra
15	36.6	9.3	397	4 ABS32578	Abs32578 Human liv
16	36.6	9.3	397	6 ABS07656	Abs07656 Human gen
17	35.6	9.1	711	2 AAT67444	Aat67444 H. pylori
18	35.6	9.1	723	2 AAT67946	Aat67946 H. pylori
19	34.8	8.9	110000	5 ABL17172	Continuation (3 of
20	34.8	8.9	4223	4 ABL17172	Continuation (3 of
21	34.4	8.8	14041	4 AAH48024	AAH48024 Internal

C 22	34.4	8.8	83391	6 ABQ67094	Abq67094 Human ang
C 23	34	8.7	110000	2 AAX91990_11	Continuation (12 o
C 24	33.4	8.5	2050	6 ABK15056	Abk15056 A. thalia
C 25	33.4	8.5	200000	12 ADO47192	Ado47192 DNA sequ
C 26	33	8.4	500	6 ABQ39467	Abq39467 Oligonuc
C 27	33	8.4	500	6 ABQ39466	Abq39466 Oligonuc
C 28	33	8.4	2211	12 ADM67052	Adm67052 Murine ad
C 29	33	8.4	4925	8 ADA22696	Ada22696 A. gossyp
C 30	33	8.4	8876	2 AAV52303	Aav52303 Streptoco
C 31	33	8.4	110000	10 ABS56454_07	Continuation (8 of
C 32	32.8	8.3	911	3 AAC38627	Aac38627 Arabidops
C 33	32.8	8.3	6476	4 ABL03322	Ab103322 Drosophil
C 34	32.6	8.3	234	6 ABK79579	Abk79579 Bacillus
C 35	32.6	8.3	415	6 ABL79948	Ab179948 Human Ova
C 36	32.4	8.2	289	4 AAL09331	Aal09331 Human bre
C 37	32.4	8.2	2285	10 ADA53466	Ada53466 Human cod
C 38	32.2	8.2	813	6 ABO46886	Abq46886 Oligonuc
C 39	32.2	8.2	813	6 ABO46887	Abq46887 Oligonuc
C 40	32.2	8.2	1989	8 ACA30799	Aca30799 Prokaryot
C 41	32.2	8.2	3128	8 ABZ10017	Abz10017 Haematopo
C 42	32.2	8.2	3128	8 ABZ10163	Abz10163 Haematopo
C 43	32.2	8.2	6028	10 ADB84205	Ades84205 Human lym
C 44	32.2	8.2	6028	10 ADB84129	Ades84129 Human lym
C 45	32.2	8.2	16648	4 ABL14476	Ab114476 Drosophil

ALIGNMENTS

RESULT 1
ABD00692/c
ID ABD00692 standard; DNA; 258 BP.
XX
AC ABD00692;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polynucleotide seqid 6467.
XX
KW Recombinant expression vector; transcription regulatory element;
XX Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.
XX
DR P-PSDB; ABO67121.
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
PS Disclosure; SEQ ID NO 6467; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX pneumoniae polypeptide of the invention
XX
SQ Sequence 258 BP; 40 A; 66 C; 53 G; 99 T; 0 U; 0 Other;

CC reactivity and reproducibility); also synthetic probes are generally
 CC shorter than probes prepared by polymerase chain reaction. ACD69191 to
 CC ACD81540 represent oligonucleotide probes used with the biochip described
 CC in the invention

XX Sequence 100 BP; 16 A; 24 C; 23 G; 37 T; 0 U; 0 Other;

Query Match 25.4%; Score 100; DB 8; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.9e-21;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TCAACGACGACGTCGCGATATCAATAAAACAAGCAGCAATAGAGCGCGGTGTGAACAA 67
 Db 258 TCAACGACGACGTCGCGATATCAATAAAACAAGCAGCAATAGAGCGCGGTGTGAACAG 199
 QY 68 CCGCGATGCCAGACAAAGTGTGATACCTTGTGGCAAGTACCGCGCCACCGCCA 127
 Db 198 TGCTGGTGCCAGACAAATCGTAATCGCGTCGGAACCGACGAGTCGGTGCACGAGCA 139
 QY 128 GAAATCGCAGTATGATAAAAGCAATCAAAAATATCGTGGGAGGAGCAGCCCAACAC 187
 Db 138 AAGATAGCCAGCAGCAAAAGCACTCGAGAGATCGCGGAACGATAGCGGAAAT 79
 QY 188 CAGGATACCGGAAACAGACTCATGTTGACCTTGTGTTAAAGAG 232
 Db 78 CAGGATACCGGAAACAGACTCATGATGACCTTGAATAAGAGCG 34

RESULT 2
 ACD69190/c
 ID ACD69190 standard; DNA; 100 BP.
 AC ACD69190;
 XX ACD69190;
 DT 18-SEP-2003 (first entry)
 DE E. coli K12 MG1655 biochip probe SEQ ID 460.
 XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.
 XX Escherichia coli.
 OS Escherichia coli.
 XX EPI260592-A1.
 XX 27-NOV-2002.
 XX 17-MAY-2001; 2001EP-00112179.
 XX 17-MAY-2001; 2001EP-00112179.
 XX 17-MAY-2001; 2001EP-00112179.
 XX (MWGB-) MWG-BIOTECH AG.
 XX Donner H, Drescher B, Huber A, Weber J;
 XX WPI; 2003-241155/24.
 XX Biochip containing probes complementary with open reading frames in
 PT Escherichia coli K12, useful for detecting gene expression and expression
 PT patterns.
 PS Claim 3; Page 82; 2004pp; German.

CC This invention describes a novel biochip comprising probe spots, each
 CC containing many identical probes. The probes are nucleotide sequences of
 CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
 CC least one includes a segment of at least 20 bases identical with, or
 CC complementary to, a segment of an open reading frame (orf) of Escherichia
 CC coli K12. The biochip is used for specific detection of gene expression
 CC in K12 and for determining the gene expression pattern, e.g. for
 CC diagnostic determination of which E. coli strains are present in the gut,
 CC and to determine the effects of e.g. growth media on gene expression. The
 CC biochip provides as comprehensive as possible detection of the K12
 CC genome, with simultaneous analysis of many different genes with a single
 CC device, and comparison of gene expression between K12 and its mutants or
 CC other E. coli strains in a single experiment. Apart from qualitative and
 CC quantitative information about gene expression, it also allows
 CC measurements of population densities for the various strains. The use of
 CC synthetic oligonucleotides for preparation of probes allows free
 CC variation in probe length and ensures high purity (and thus selectivity,
 CC reactivity and reproducibility); also synthetic probes are generally
 CC shorter than probes prepared by polymerase chain reaction. ACD69191 to

CC reactivity and reproducibility); also synthetic probes are generally
 CC shorter than probes prepared by polymerase chain reaction. ACD69191 to
 CC ACD81540 represent oligonucleotide probes used with the biochip described
 CC in the invention

XX Sequence 100 BP; 16 A; 24 C; 23 G; 37 T; 0 U; 0 Other;

Query Match 25.4%; Score 100; DB 8; Length 100;
 Best Local Similarity 100.0%; Pred. No. 2.9e-21;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TCAACGACGACGTCGCGATATCAATAAAACAAGCAGCAATAGAGCGCGGTGTGAACAA 67
 Db 100 TCAACGACGACGTCGCGATATCAATAAAACAAGCAGCAATAGAGCGCGGTGTGAACAA 41
 QY 68 CCGCGATGCCAGACAAAGTGTGATACCTTGTGGCACA 107
 Db 40 CCGCGATGCCAGACAAAGTGTGATACCTTGTGGCACA 1

RESULT 3
 ACD69191/c
 ID ACD69191 standard; DNA; 100 BP.
 AC ACD69191;
 XX ACD69191;
 DT 18-SEP-2003 (first entry)
 DE E. coli K12 MG1655 biochip probe SEQ ID 461.
 XX Biochip; gene expression; gut; diagnostic; detection; probe; ss.
 XX Escherichia coli.
 OS Escherichia coli.
 XX EPI260592-A1.
 XX 27-NOV-2002.
 XX 17-MAY-2001; 2001EP-00112179.
 XX 17-MAY-2001; 2001EP-00112179.
 XX 17-MAY-2001; 2001EP-00112179.
 XX (MWGB-) MWG-BIOTECH AG.
 XX Donner H, Drescher B, Huber A, Weber J;
 XX WPI; 2003-241155/24.
 XX Biochip containing probes complementary with open reading frames in
 PT Escherichia coli K12, useful for detecting gene expression and expression
 PT patterns.
 PS Claim 3; Page 82; 2004pp; German.

CC This invention describes a novel biochip comprising probe spots, each
 CC containing many identical probes. The probes are nucleotide sequences of
 CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
 CC least one includes a segment of at least 20 bases identical with, or
 CC complementary to, a segment of an open reading frame (orf) of Escherichia
 CC coli K12. The biochip is used for specific detection of gene expression
 CC in K12 and for determining the gene expression pattern, e.g. for
 CC diagnostic determination of which E. coli strains are present in the gut,
 CC and to determine the effects of e.g. growth media on gene expression. The
 CC biochip provides as comprehensive as possible detection of the K12
 CC genome, with simultaneous analysis of many different genes with a single
 CC device, and comparison of gene expression between K12 and its mutants or
 CC other E. coli strains in a single experiment. Apart from qualitative and
 CC quantitative information about gene expression, it also allows
 CC measurements of population densities for the various strains. The use of
 CC synthetic oligonucleotides for preparation of probes allows free
 CC variation in probe length and ensures high purity (and thus selectivity,
 CC reactivity and reproducibility); also synthetic probes are generally
 CC shorter than probes prepared by polymerase chain reaction. ACD69191 to

CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention

XX SQ Sequence 100 BP; 13 A; 22 C; 27 G; 38 T; 0 U; 0 Other;
Query Match 25.4%; Score 100; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.9e-21;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 CTGTGTCACAGTACCGGCGGACAGCCAGCAAGAAATCGCAGTGAATAAGCAATTCA 155
DB 100 CTGTGTCACAGTACCGGCGGACAGCCAGCAAGAAATCGCAGTGAATAAGCAATTCA 41
QY 156 AAAAATATCGGTGGGAAGACAGCCCAACACACAGATAA 195
DB 40 AAAATATCGGTGGGAAGACAGCCCAACACACAGATAA 1

RESULT 4
ACD69192/C
ID ACD69192 standard; DNA; 100 BP.
XX
AC ACD69192;
XX
DT 18-SEP-2003 (first entry)
XX
DE E. coli K12 M61655 biochip probe SEQ ID 462.
XX
KW Biochip; gene expression; gut; diagnostic; detection; probe; ss.
XX
OS Escherichia coli.
XX
PN EP1260592-A1.
XX
PD 27-NOV-2002.
XX
PF 17-MAY-2001; 2001EP-00112179.
XX
PR 17-MAY-2001; 2001EP-00112179.
XX
PA (MWGB-) MWG-BIOTECH AG.
XX
PI Donner H, Drescher B, Huber A, Weber J;
XX
DR WPI; 2003-241155/24.
XX
PT Biochip containing probes complementary with open reading frames in
PT Escherichia coli K12, useful for detecting gene expression and expression
PT patterns.
XX
PS Claim 3; Page 82; 2004pp; German.

XX This invention describes a novel biochip comprising probe spots, each
CC containing many identical probes. The probes are nucleotide sequences of
CC 30-80 bases, are prepared ex situ from synthetic oligonucleotides and at
CC least one includes a segment of at least 20 bases identical with, or
CC complementary to, a segment of an open reading frame (orf) of Escherichia
CC coli K12. The biochip is used for specific detection of gene expression
CC in K12 and for determining the gene expression pattern, e.g. for
CC diagnostic determination of which E. coli strains are present in the gut,
CC and to determine the effects of e.g. growth media on gene expression. The
CC biochip provides as comprehensive as possible detection of the K12
CC genome, with simultaneous analysis of many different genes with a single
CC device, and comparison of gene expression between K12 and its mutants or
CC other E. coli strains in a single experiment. Apart from qualitative or
CC quantitative information about gene expression, it also allows
CC measurements of population densities for the various strains. The use of
CC synthetic oligonucleotides for preparation of probes allows free
CC variation in probe length and ensures high purity (and thus selectivity,
CC reactivity and reproducibility); also synthetic probes are generally
CC shorter than probes prepared by polymerase chain reaction. ACD68731 to
CC ACD81540 represent oligonucleotide probes used with the biochip described
CC in the invention

XX SQ Sequence 100 BP; 23 A; 26 C; 20 G; 31 T; 0 U; 0 Other;
Query Match 25.4%; Score 100; DB 8; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.9e-21;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 161 TATCGGTGGGAAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTGACCTT 220
DB 100 TATCGGTGGGAAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTGACCTT 41
QY 221 GGTTGTAAGAGAGAGCAGCGGCTTATTATTTCAGCATCT 260
DB 40 GGTTGTAAGAGAGAGCAGCGGCTTATTATTTCAGCATCT 1

RESULT 5
AAV12742
ID AAV12742 standard; cDNA to mRNA; 1181 BP.
XX
AC AAV12742;
XX
DT 05-MAY-1998 (first entry)
XX
DE RBE1 transcription controlling factor #33 (RBEF1) coding sequence.
XX
KW RBE1; rice starch branching enzyme; transcription controlling factor; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 55..852
FT /*tag= a
XX
PN JPI0014575-A.
XX
PD 20-JAN-1998.
XX
PF 28-JUN-1996; 96JP-00169246.
XX
PR 28-JUN-1996; 96JP-00169246.
XX
PA (MITS-) MITSUI GYOSAI SHOKUBUTSU BIO KENKYUSHO.
PA (MITK) MITSUI TOATSU CHEM INC.
XX
DR WPI; 1998-138237/13.
XX
P-PSDB; AAW41152.
XX
PT Transcription controlling factor in rice - comprises starch branching
PT enzyme gene, allows rice to be improved in quality and quantity.
XX
PS Claim 3; Page 6-7; 11pp; Japanese.

XX This sequence represents the DNA encoding a transcription controlling
CC factor of the invention, designated #33 (RBEF1). The transcription
CC controlling factor in rice promotes transcription of the RBE1 gene. Use
CC of the transcription factor enables rice to be improved in quality and
CC quantity
XX
SQ Sequence 1181 BP; 315 A; 257 C; 337 G; 272 T; 0 U; 0 Other;
Query Match 9.8%; Score 38.4; DB 2; Length 1181;
Best Local Similarity 56.2%; Pred. No. 0.21;
Matches 72; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 126 CAGAAATCGCCAGTGTATTAAGCAATTCAGAAATATCGGTGGGAAGACAGCCCAAC 185
DB 778 CACAAGACACTAATAACAGCAACGGGAAGTATGTTAGTGGCAGCACAACACAGAG 837
QY 186 ACCACGATACGGGAACAGACTCATGTTGACCTTGGTTAAAGAGAGAGAGCGGTGA 245
DB 838 GCTAACTCTAGGTGAGAGAGATCATGAGCAAGTGCAGCAAGATGAGATGAGTGTTC 837

QY 246 TTATTTC 253
 Db 898 TTCTTATC 905

RESULT 6
 ACH95383/C
 ID ACH95383 standard; DNA; 1035 BP.
 XX
 AC ACH95383;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polynucleotide seqid 1178.
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.
 XX
 OS Klebsiella pneumoniae.
 PN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 XX
 DR WPI; 2003-895346/82.
 DR P-PSDB; ABO61832.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 1178; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
 CC pneumoniae polypeptide of the invention
 XX
 SQ Sequence 1035 BP; 151 A; 330 C; 312 G; 242 T; 0 U; 0 Other;
 Query Match 9.5%; Score 37.2; DB 11; Length 1035;
 Best Local Similarity 50.0%; Pred. No. 0.48;
 Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 63 AACACGCCGATGCCAGCAAGTCGTAGATACCTGTGGCACAAGTACCGCGGCACC 122
 Db 731 AACGCTCGCGCCGCCGACACTAGTCGTAGTTCGGGTACGGAGAAATCCGCGGCACC 672

QY 123 AGCCAGAAATCCGCAAGTGTATAAAGCAATTCAAAAATATCGGTGGGAAGACAGCCCA 182
 Db 671 AGCCCAACGAGGTCATCCAGGCAAGAGACGGTGATTGCCAGAGCCACCAGAGGCC 612

QY 183 AACACCAACCAACGCGGAACAGACTCATGTGTGACCTGGTTGTAAAGAGAGAGCAGCGC 242
 Db 611 GGTGTCCACCGCTGTGAGAGCAGGATATCAGGAACAGCGTTGGCATCCGCGCACGACT 552

QY 243 TTATTA 248
 Db 551 TCGATA 546

RESULT 7
 ABL34071

ID ABL34071 standard; DNA; 35000 BP.
 AC ABL34071;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Human RECQL gene-related nucleotide #1.
 KW Human; antisense therapy; ds; RECQL; hyperproliferative disorder; cancer;
 KW premature ageing; infection; inflammation; tumour formation.
 XX
 OS Homo sapiens.
 PN WO200268590-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 21-FEB-2002; 2002WO-US005225.
 XX
 PR 23-FEB-2001; 2001US-00793807.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Ward DT, Watt AT;
 XX
 DR WPI; 2002-750415/81.
 XX
 PT New antisense oligonucleotides targeted to a nucleic acid encoding RECQL,
 PT useful for modulating the expression of RECQL protein, or for treating a
 PT disease or condition associated with the expression of RECQL, e.g.
 PT cancer.
 XX
 PS Example 15; Page 101-121; 138pp; English.
 XX
 CC The invention comprises antisense oligonucleotides which inhibit
 CC expression of the human RECQL gene. The antisense oligonucleotides of the
 CC invention are useful for modulating the expression of RECQL protein and
 CC in treating hyperproliferative disorders (e.g. cancer and conditions
 CC involving premature ageing). The antisense oligonucleotides of the
 CC invention are also useful for diagnostics, therapeutics and prophylaxis
 CC (e.g. to prevent or delay infection, inflammation or tumour formation).
 CC The present DNA sequence was used in the invention
 XX
 SQ Sequence 35000 BP; 10326 A; 6524 C; 6774 G; 11376 T; 0 U; 0 Other;
 Query Match 9.4%; Score 37; DB 6; Length 35000;
 Best Local Similarity 49.2%; Pred. No. 2.6;
 Matches 97; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 192 ATAACGGGAACAGACTCATGTGTGACCTGGTTGTAAAGAGAGAGCAGCGCTATTATT 251
 Db 3747 ATTAGTGTCAAGAACACCATGTGTTCTTATTGAAAGTCACTGCAATATATATTAC 3806

QY 252 TCAGCATCTGTCCCGCAGAGAGGCGCATGGAAAGCCGGCGGAGAGCAACATTCGTGTAG 311
 Db 3807 TTCTTAACATTTCCCGTGTGGATGACAGATTACTGTTGAGTGACATCATTAATAG 3866

QY 312 ATTGATATTAAATATATTAGCGTAAGTGTATGCTGTATCTATATATTATGTGATCTAAAT 371
 Db 3867 TACTATTTTAAAAATTTGTTGTCATGTTGCTGCTTTTATTTTAGAAGAAATTTGGAT 3926

QY 372 CACTTTTAAAGTCAGACT 388
 Db 3927 TATTTTGAAGCAAAAT 3943

RESULT 8
 ABL34071
 ID ABL34071 standard; DNA; 13503 BP.
 XX
 AC ABL34071;
 XX
 DT 26-MAR-2002 (first entry)

XX DE Human immune system associated gene SEQ ID NO: 2044.

XX DE Human; immune system disease; cytosine methylation; antiasthmatic;

XX DE antiarteriosclerotic; antianaemic; cytosine; cytosine; cytosine; cytosine;

XX DE neuroprotection; anti-HIV; anticonvulsant; ophthalmological;

XX DE antirheumatic; antidiabetic; antidiabetic; antipsoriasis;

XX DE antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;

XX DE acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

XX DE neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

XX OS Homo sapiens.

XX PN WO200200928-A2.

XX PN 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007537.

XX PF 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX XX (EPG-) EPIGENOMICS AG.

XX XX Olek A, Piepenbrock C, Berlin K;

XX PI WPI; 2002-130909/17.

XX DR Nucleic acid comprising fragment of chemically modified gene, useful for

XX PT diagnosis and treatment of diseases associated with abnormal cytosine

XX PT methylation.

XX XX Claim 1; SEQ ID NO 2044; 32pp + Sequence Listing; German.

XX PS The present invention provides a number of human immune system associated

XX CC genes which are modified by the methylation of cytosines. The sequences

XX CC can be used in the diagnosis and treatment of immune system disorders,

XX CC including eye diseases such as retinopathy, neovascular glaucoma and

XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

XX CC diseases. The present sequence is a gene of the invention

XX XX Sequence 13503 BP; 3398 A; 291 C; 3528 G; 6286 T; 0 U; 0 Other;

XX SQ Query Match 9.4%; Score 36.8; DB 6; Length 13503;

XX Best Local Similarity 63.6%; Pred. No. 2;

XX Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 299 ACATTCCTGTAGATGATATTAATATTAATAGCTAGCTGTTATCTATATT 358

Db 10532 AATATGATGAAAGTAGATATTTTAAATAGTGTAAATTTTATTTTAAAT 10591

QY 359 ATGTATCTAAATCACTTTTAAGTCAGA 386

Db 10592 TAGTTATATAATTAGTGAATTTTCA 10619

RESULT 9

AAI16233

ID AAI16233 standard; DNA; 397 BP.

XX AC AAI16233;

XX AC AAI16233;

XX DT 12-OCT-2001 (first entry)

XX DE Probe #6166 for gene expression analysis in human cervical cell sample.

XX DE Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW Cervical cancer; ss.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PN 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX DR Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human cervical epithelial cells.

XX PS Claim 25; SEQ ID NO 6166; 487pp; English.

XX XX The present invention relates to human single exon nucleic acid probes

XX CC (SENPs). The present sequence is one such probe. The SENPs are derived

XX CC from human HeLa cells. The SENPs can be used to produce a single exon

XX CC microarray, which can be used for measuring human gene expression in a

XX CC sample derived from human cervical epithelial cells. By measuring gene

XX CC expression, the probes are therefore useful in grading and/or staging of

XX CC diseases of the cervix, notably cervical cancer. Note: The sequence data

XX CC for this patent did not form part of the printed specification, but was

XX CC obtained in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX XX Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;

XX SQ Query Match 9.3%; Score 36.6; DB 4; Length 397;

XX Best Local Similarity 47.2%; Pred. No. 0.49;

XX Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 156 AAAAATATCGTGGGAGGACAGCCCAACACACGATACGCGGAAACAGACTCATGTTG 215

Db 135 AAAAATATCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 194

QY 216 ACCTTGCTTCTAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 275

Db 195 AACCTTCATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 254

QY 276 GGATGGAAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 335

Db 255 TGCTGATAGAGT 314

QY 336 ACTGTTATGCTGTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 390

Db 315 TAGAATGACACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369

RESULT 10

ABA58948

ID ABA58948 standard; DNA; 397 BP.

XX AC ABA58948;

XX AC ABA58948;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #7253.

XX DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.


```
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488990/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 1; SEQ ID NO 6274; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease. Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 9.3%; Score 36.6; DB 4; Length 397;
XX Best Local Similarity 47.2%; Pred. No. 0.49;
XX Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
XX
XX QY 156 AAAAATATCGTGGGAGAGACAGCCCAACACACGATACCGGAAACAGACTCATGTTG 215
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 135 AAAAATATCGTGGGAGAGACAGCCCAACACACGATACCGGAAACAGACTCATGTTG 194
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 216 ACCTTGGTTGTTAAAGAGAGAGAGCGGTTATTTTTCAGCATCTGTGCGCGCAGAGAAG 275
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 195 AACCTTCATCTAAAGTGATATAAATAGTTATTACATACATCAAGCTCCAGCCTTGATC 254
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 276 GGCATGGAAGCGGCGAGAGACACATTCGTTAGATTGATTTAATATATATAGCGTA 335
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 314
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 336 ACTGTTATGCTGTATCTATATTATGATCTAAATCACTTTTAAGTCAGAGTGA 390
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 315 TAGAGATGACACCTGGAGGAGATATAATGTTACACAAATTAAGTCATATAGA 369
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 13
XX AAK32852
XX ID AAK32852 standard; DNA; 397 BP.
XX
XX AC AAK32852;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 7409.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
```

```
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488990/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 7409; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
XX Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX
XX Query Match 9.3%; Score 36.6; DB 4; Length 397;
XX Best Local Similarity 47.2%; Pred. No. 0.49;
XX Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
XX
XX QY 156 AAAAATATCGTGGGAGAGACAGCCCAACACACGATACCGGAAACAGACTCATGTTG 215
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 135 AAAAATATCGTGGGAGAGACAGCCCAACACACGATACCGGAAACAGACTCATGTTG 194
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 216 ACCTTGGTTGTTAAAGAGAGAGAGCGGTTATTTTTCAGCATCTGTGCGCGCAGAGAAG 275
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 195 AACCTTCATCTAAAGTGATATAAATAGTTATTACATACATCAAGCTCCAGCCTTGATC 254
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 276 GGCATGGAAGCGGCGAGAGACACATTCGTTAGATTGATTTAATATATATAGCGTA 335
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 314
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 336 ACTGTTATGCTGTATCTATATTATGATCTAAATCACTTTTAAGTCAGAGTGA 390
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 315 TAGAGATGACACCTGGAGGAGATATAATGTTACACAAATTAAGTCATATAGA 369
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 14
XX AAK07110
XX ID AAK07110 standard; DNA; 397 BP.
XX
XX AC AAK07110;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe SEQ ID NO: 7101.
XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
```

```
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-488898/53.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX PS Example 4; SEQ ID NO 7101; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention
XX CC
XX CC Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX CC
XX CC Query Match 9.3%; Score 36.6; DB 4; Length 397;
XX CC Best Local Similarity 47.2%; Pred. No. 0.49;
XX CC Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 156 AAAAAATATCGTGGGAGAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTG 215
DB 135 AAAAAATATGTTTAAAGAAATGAAAGATCTGGCTCTTGGTCAACAGGCAAAAGACA 194
QY 216 ACCTTGGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 275
DB 195 AACCTTCATCTAAAGTGATATAAATAGTATTACATACATCATAGCTCCAGCCTTGATC 254
QY 276 GGATGGAAAGCCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 335
DB 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAGAGGTTGGATACCCCAAGTAGAATA 314
QY 336 ACTGTTATCTGTATCTATATATGATCTAAATCACTTTTAAGTCAGAGTGA 390
DB 315 TAGAGATGACACCTGGAGGAGATATAATGTTAAACAAATTTAAGTCATATAGA 369
RESULT 15
ABS32578
ID ABS32578 standard; DNA; 397 BP.
XX AC ABS32578;
XX XX
XX DT 25-FEB-2003 (first entry)
XX DE
XX DE Human liver single exon probe, SEQ ID NO 7568.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX XX
```

```
PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX XX WPI; 2001-488898/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 1; SEQ ID NO 7568; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABS25011-ABS51005 represent human
XX CC liver single exon nucleic acid probes of the invention. Note: The
XX CC sequence information for this patent does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX CC
XX CC Sequence 397 BP; 145 A; 66 C; 78 G; 108 T; 0 U; 0 Other;
XX CC
XX CC Query Match 9.3%; Score 36.6; DB 4; Length 397;
XX CC Best Local Similarity 47.2%; Pred. No. 0.49;
XX CC Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 156 AAAAAATATCGTGGGAGAGACAGCCCAACACACAGATAACGGGAACAGACTCATGTTG 215
DB 135 AAAAAATATGTTTAAAGAAATGAAAGATCTGGCTCTTGGTCAACAGGCAAAAGACA 194
QY 216 ACCTTGGTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 275
DB 195 AACCTTCATCTAAAGTGATATAAATAGTATTACATACATCATAGCTCCAGCCTTGATC 254
QY 276 GGATGGAAAGCCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 335
DB 255 TGCTGATAGAGTCTGTGTTGGTCTTCATCCAGAGAGGTTGGATACCCCAAGTAGAATA 314
QY 336 ACTGTTATCTGTATCTATATATGATCTAAATCACTTTTAAGTCAGAGTGA 390
DB 315 TAGAGATGACACCTGGAGGAGATATAATGTTAAACAAATTTAAGTCATATAGA 369
Search completed: December 18, 2004, 23:21:43
Job time : 261.841 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 15:34:11 ; Search time 43.9637 Seconds
(without alignments)
6353.876 Million cell updates/sec

Title: US-10-759-889-3
Perfect score: 393
Sequence: 1 tccaacctcaaacgacagt.....ctttaagtcagagtgaata 393
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	146.6	37.3	258	US-09-489-039A-6467	Sequence 6467, Ap
C 2	44.2	11.2	7218	US-08-232-463-14	Sequence 14, Appl
C 3	37.2	9.5	1035	US-09-489-039A-1178	Sequence 1178, Ap
C 4	34	8.7	1230025	US-09-198-452A-1	Sequence 1, Appli
C 5	33	8.4	8876	US-08-961-527-170	Sequence 170, App
C 6	32	8.1	240	US-09-252-991A-3764	Sequence 3764, Ap
C 7	32	8.1	1383	US-09-252-991A-3609	Sequence 3609, Ap
C 8	31.6	8.0	2473	US-09-173-914-3	Sequence 3, Appli
C 9	31.4	8.0	3768	US-09-248-796A-55	Sequence 55, Appl
C 10	31	7.9	2289	US-09-614-221A-375	Sequence 375, App
C 11	30.8	7.8	369	US-09-270-767-1375	Sequence 1375, App
C 12	30.8	7.8	369	US-09-270-767-17257	Sequence 17257, A
C 13	30.6	7.8	336	US-09-060-756-250	Sequence 250, App
C 14	30.6	7.8	336	US-09-670-314-250	Sequence 250, App
C 15	30.6	7.8	375	US-09-621-976-8473	Sequence 8473, Ap
C 16	30.6	7.8	118067	US-09-621-976-8473	Sequence 8473, Ap
C 17	30.6	7.8	118067	US-09-621-976-8473	Sequence 8473, Ap
C 18	30.4	7.7	852	US-09-134-000C-3192	Sequence 3192, Ap
C 19	30.2	7.7	403	US-09-252-991A-7045	Sequence 7045, Ap
C 20	30.2	7.7	450	US-09-513-993C-31966	Sequence 31966, A
C 21	30.2	7.7	1284	US-09-107-532A-2979	Sequence 2979, Ap
C 22	30	7.6	736	US-09-270-767-7757	Sequence 7757, Ap
C 23	30	7.6	736	US-09-270-767-23039	Sequence 23039, A
C 24	30	7.6	1001	US-09-671-317-364	Sequence 364, App
C 25	30	7.6	1001	US-09-671-317-365	Sequence 365, App
C 26	30	7.6	1001	US-09-671-317-366	Sequence 366, App
C 27	30	7.6	1001	US-09-671-317-367	Sequence 367, App

C 28	30	7.6	168575	4	US-09-426-290-1	Sequence 1, Appli
C 29	29.8	7.6	3026	4	US-09-963-137-136	Sequence 136, App
C 30	29.8	7.6	3026	4	US-09-963-137-140	Sequence 140, App
C 31	29.8	7.6	3172	4	US-09-976-594-940	Sequence 940, App
C 32	29.8	7.6	6423	4	US-09-774-528-244	Sequence 244, App
C 33	29.8	7.6	33769	4	US-09-544-398B-8	Sequence 8, Appli
C 34	29.8	7.6	33769	4	US-09-543-771-8	Sequence 8, Appli
C 35	29.6	7.5	193303	4	US-09-497-855A-37	Sequence 37, Appl
C 36	29.6	7.5	193303	4	US-09-497-855A-44	Sequence 44, Appl
C 37	29.4	7.5	427	4	US-09-621-976-17978	Sequence 17978, A
C 38	29.4	7.5	474	4	US-09-621-976-18033	Sequence 18033, A
C 39	29.4	7.5	581	4	US-09-270-767-10440	Sequence 10440, A
C 40	29.4	7.5	613	2	US-08-740-309-1	Sequence 1, Appli
C 41	29.4	7.5	1200	4	US-09-252-991A-7018	Sequence 7018, Ap
C 42	29.4	7.5	1568	4	US-09-799-451-653	Sequence 653, App
C 43	29.4	7.5	1587	4	US-09-489-039A-6670	Sequence 6670, Ap
C 44	29.4	7.5	2469	4	US-09-489-039A-6749	Sequence 6749, Ap
C 45	29.4	7.5	22846	2	US-08-469-461-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-489-039A-6467/c
; Sequence 6467, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6467
; LENGTH: 258
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6467

Query Match	37.3%	Score	146.6	DB	4	Length	258
Best Local Similarity	78.2%	Pred. No.	4.4e-39				
Matches	176	Conservative	0	Mismatches	49	Indels	0
Gaps	0						
Qy	8	TCAAACGAACGATCGGATATCAATAAAACAAAGCAGCAATAGAGCGCGTGTGAACAA	67				
Db	258	TCAAACGAACGATCGGATATCAATAAAACAAAGCAGCAATAGAGCGCGTGTGAACAA	199				
Qy	68	CGCGGATGCCAGACAAAGTCGTAGATACCTGTGGCACAAGTACCGCGCGCAGCCCA	127				
Db	198	TGCTGGTGCAGACAAATCGTAAATGCGGTGCGAACCCAGCATCGGTGACCAAGCA	139				
Qy	128	GAAAAATCGCAGTGATAAAGCAATTCAAAAATATCGTGGGAAGGAGAGCCCAACAC	187				
Db	138	AAAGATAGCCAGCAGCAAAAGCAACTCTGAAGAAGATCGCGGGAACGATTAAGCGAAAT	79				
Qy	188	CACGATAACGGAAACAGACTCATGTGACCTTGTGTTAAAGAG	232				
Db	78	CACGATGACAGGAACAGACTCATGTGACCTTGTGTTAAAGAG	34				

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMQU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9pt-Fls
US-08-232-463-14

Query Match 11.2%; Score 44.2; DB 1; Length 7218;
Best Local Similarity 1.3%; Pred. No. 0.00035;
Matches 4; Conservative 180; Mismatches 113; Indels 0; Gaps 0;
QY 10 AACGACAGTCGCGATATCAATAAAACAGCAGCAATAGACGCGGTGTGAACAACG 69
DB 1358 RRR 1299
QY 70 CCGATGCCAGCAAGTCGTAGATACCTGTTGGCACAAGTACCCGCGCCACGACCA 129
DB 1298 RRR 1239
QY 130 AATCCCGAGTATAAAGCAATTCATAAAATATCGTGGGAGGACGACCCCAACCA 189
DB 1238 RRR 1179
QY 190 CGATAACGGGAACAGACACTCATGTCACCTTGTGTAAGAGAGAGACGCGTTATTAT 249
DB 1178 RRR 1119
QY 250 TTTCAGCATCTGTCGGCGAGAGAGCGGATGGAAGCCGCGGAGAGCAACATTC 306
DB 1118 RRR 1062

RESULT 3
US-09-489-039A-1178/c
; Sequence 1178, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.20040001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 1178
LENGTH: 1035
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1178
Query Match 9.5%; Score 37.2; DB 4; Length 1035;
Best Local Similarity 50.0%; Pred. No. 0.026;
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 63 AACACGCCCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGCGGACCC 122
DB 731 AACGCTGGCGCCGCCGCAACATAGTCGTAGTTCGGGTACGGAGAAATTCGCGCGGACCC 672
QY 123 AGCCAGAAATCCCGAGTGTATAAAGCAATTCAAAAATATCGTGGGAGGACGACCCCA 182
DB 671 ACGCCACACGCGTCATCCAGCCAAAGAGAACGTCGATTCGCCAGACGCCACGAGAGCCC 612
QY 183 AACACACGATACACGCGGAAACAGACTCATGTTGACCTTGTGTAAGAGAGAGAGCGGCG 242
DB 611 GGCTGCACGACGCTGGAGAGCAGGATAATCAGGAACAGCGTTGGCATGCCGCGGACGACT 552
QY 243 TTATTA 248
DB 551 TCGATA 546

RESULT 4
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature


```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; FILE REFERENCE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; CURRENT FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3609
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3609

Query Match
Best Local Similarity 8.1%; Score 32; DB 4; Length 1383;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 37 ACAAGCAGCAATAGAGCGCGGTGTTGAACAAGCCCGGATGCCAGACAAGTCTAGATAC 96
Db 1137 ACAGGCGAGCAACAGGCTGACCCGGAACACAGCGCGGTGCCAGGTGAAGCATACAGGC 1196

QY 97 CTGTTGGCACAAGTACCCCGGCGCACCAGCAGCAAAATGCCAGTGTATAAAGCAATTCAA 156
Db 1197 CAACGAGCGAAGATCCGGTCCGAGCCCCCAGGTGAGCGCCGCGCATCACGAACAACA 1256

QY 157 AAAATATCGGTGGGAGGACAGCCCAACACACCGATACACGGGAACA 204
Db 1257 GCGTCAGGTCGGCATGTAGACCCCATGATGCGCATTCACGAGGCA 1304

RESULT 8
US-09-173-914-3
; Sequence 3, Application US/09173914
; Patent No. 6171857
; GENERAL INFORMATION:
; APPLICANT: Hendrickson, Eric
; TITLE OF INVENTION: A No. 6171857el Leucine Zipper, KARP-1 and
; FILE REFERENCE: B0877/7017/HK
; CURRENT APPLICATION NUMBER: US/09/173,914
; CURRENT FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 60/064,557
; EARLIER FILING DATE: 1997-10-17
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2473
; TYPE: DNA
; ORGANISM: Mesocricetus Auratus
US-09-173-914-3

Query Match
Best Local Similarity 8.0%; Score 31.6; DB 3; Length 2473;
Matches 79; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 155 AAAAATATCGTGGGAGGACAGCCCAACACACAGATACGGGAACAGACTCATGTT 214
Db 1673 AAAAAGGATCAAGTACCTGCTCAGGACATTTTCCAAGACAAATGATGAAGGGGCGCTGCT 1732

QY 215 CACCTTGTTGTTAAAGAGAGACAGCGCTTATTTTTCAGCATCTGTCGCGCGAGAGAA 274
Db 1733 GCCAAAAGTGTAAATGGAGAAGAGAGATCATCATCATCTCTAGCCTGGCTGAA 1792

QY 275 GGGCATGAAGCCGGGCGAGGCAACATTTGCTGTAGA 312
Db 1793 GGGAAATGTACACAGGTTTGAAGTGTGAATCTCTGTGA 1830

RESULT 9
US-09-248-796A-55/c
; Sequence 55, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 55
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-55

Query Match
Best Local Similarity 8.0%; Score 31.4; DB 4; Length 3768;
Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 302 ATTGCTGTAGATTGATATTTTAATATATATACGTAACCTGTATCTGTATCTATATTATG 361
Db 3137 AATGATCCAAATTTGATATTTTATTTTCATATTTTAAATCACTTAAATTTAATATA 3078

QY 362 TGATCTAATCACTTTTAAGTCAGAGTGA 390
Db 3077 GGATCTAAATCTGGTGTCTCTGATTGTA 3049

RESULT 10
US-09-614-221A-375/c
; Sequence 375, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunanandaa, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; TITLE OF INVENTION: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 375
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-375

Query Match
Best Local Similarity 7.9%; Score 31; DB 4; Length 2289;
Matches 76; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 134 CGCCAGTGTATAAAGCAATTCAAAAAATATCGTGGGAAGGACAGCCCAACACACCGAT 193
Db 2211 CACTAGTGGTGAAGATATACAAAAAACCCGCAACAGCACAGGAGACACACACCAAGAA 2152

QY 194 AACGGAAACAGACTCATGTTGACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTTC 253
Db 2151 GACGGTATAGGTTTAACTAGACCTTAGGTTCCGTTTCATATGATGACCTGAAATATC 2092

QY 254 AGCATCTGTCGCGCAGAGAAGGGCATGAA 284
Db 2091 CTCGTTTCTCCAAATCCATGATTGCA 2061

RESULT 11
US-09-270-767-1975
; Sequence 1975, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
```


Search completed: December 18, 2004, 20:49:00
Job time : 48.9637 secs

This Page Blank (uspto)

GenCore version 5.1.6		Copyright (c) 1993 - 2004 Compugen Ltd.	
OM nucleic - nucleic search, using sw model			
Run on:	December 18, 2004, 20:42:54	Search time 272.694 Seconds	
		(without alignments)	
		7956.501 Million cell updates/sec	
Title:	US-10-759-889-3		
Perfect score:	393		
Sequence:	1 ttcaactcacaacgaacagt.....ctttaagtgcagtggaata 393		
Scoring table:	IDENTITY NUC		
	Gapop 10.0, Gapext 1.0		
Searched:	4093002 seqs, 2760418925 residues		
Total number of hits satisfying chosen parameters:	8186004		
Minimum DB seq length:	0		
Maximum DB seq length:	2000000000		
Post-processing:	Minimum Match 0%		
	Maximum Match 100%		
	Listing first 45 summaries		
Database :	Published Applications NA:*		
	1: /cgn2_6/ptodata/2/pubnra/US07 PUBCOMB.seq:*		
	2: /cgn2_6/ptodata/2/pubnra/PCT_NEW_PUB.seq:*		
	3: /cgn2_6/ptodata/2/pubnra/US06_NEW_PUB.seq:*		

ALIGNMENTS

```

RESULT 1
US-10-759-889-3
; Sequence 3, Application US/10759889
; Publication No. US20040157331A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Co., Inc.
; APPLICANT: Van Dyk, Tina K
; TITLE OF INVENTION: Regulator/Promoter for Tunable Gene Expression and Metabolite
; TITLE OF INVENTION: Sensing
; FILE REFERENCE: CL2036 US NA
; CURRENT APPLICATION NUMBER: US/10/759,889
; CURRENT FILING DATE: 2004-01-16
; PRIOR APPLICATION NUMBER: 60/440,965
; PRIOR FILING DATE: 2003-01-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-759-889-3

```

Result No.	Score	Match	Query Length	DB ID	Description
1	393	100.0	393	17	US-10-759-889-3
2	39	9.9	1842	18	US-10-425-115-137044
3	38.4	9.8	1598	17	US-10-437-963-70841
4	37	9.4	3500	10	US-09-793-807-10
5	36.8	9.4	13503	15	US-10-311-455-2044
6	36.6	9.3	397	9	US-09-864-761-6274
7	36.6	9.3	330973	12	US-10-087-192-1498
8	35.6	9.1	711	16	US-10-335-977-4100
9	35.6	9.1	714	16	US-10-335-977-4099
10	35.6	9.1	723	16	US-10-335-977-4101
11	35	8.9	513509	10	US-09-754-853A-4
12	34.4	8.8	83391	17	US-10-433-793-124

Db 121 CCAGCCAGAAATCCGAGTGAATAAGCAATTCAAAATAATCGGTGGGAAGACAGCC 180
QY 181 CAAACACCAAGTACGAGGAAACAGACTCATGTTGACCTTGGTCTAAAGAGAGAGCAGG 240
Db 181 CAAACACCAAGTACGAGGAAACAGACTCATGTTGACCTTGGTCTAAAGAGAGAGCAGG 240
QY 241 CGTTATTAATTTTTCAGACTCTGTCGCGCAGAGAGAGGCGATGGAAGCCGGCGAGAGCAA 300
Db 241 CGTTATTAATTTTTCAGACTCTGTCGCGCAGAGAGAGGCGATGGAAGCCGGCGAGAGCAA 300
QY 301 CATTCTGTAGATTGATATTTAATATATATAGGTAACGTTATGCTGTATCTATATAT 360
Db 301 CATTCTGTAGATTGATATTTAATATATATAGGTAACGTTATGCTGTATCTATATAT 360
QY 361 GTGATCTAATCACTTTTAAAGTCAGAGTGAATA 393
Db 361 GTGATCTAATCACTTTTAAAGTCAGAGTGAATA 393

RESULT 2

US-10-425-115-137044
; Sequence 137044, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 137044
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_56460C.1
US-10-425-115-137044

Query Match 9.9%; Score 39; DB 18; Length 1842;
Best Local Similarity 58.0%; Pred. No. 0.41; Mismatches 50; Indels 0; Gaps 0;
Matches 69; Conservative 0;

QY 67 ACGCCGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACAGCC 126
Db 1574 ACAACAGTACCCAAACTAGCTCAAACTTCTTGTGGCCTAAACTAAGAGATCCTCCAGCC 1633
QY 127 AGAAATCCCGAGTGAATAAGCAATTCAAAATAATCGGTGGGAGGACAGACCCAAAC 185
Db 1634 CTTAGTATCCCGAGGAGATCAACACAAACATATCTGTACCAAGGAGCTCCATGAC 1692

RESULT 3

US-10-437-963-70841
; Sequence 70841, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 70841
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71373C.1
US-10-437-963-70841

Query Match 9.8%; Score 38.4; DB 17; Length 1598;
Best Local Similarity 56.2%; Pred. No. 0.58; Mismatches 56; Indels 0; Gaps 0;
Matches 72; Conservative 0;

QY 126 CAGAAATCCGAGTGAATAAGCAATTCAAAATAATCGGTGGGAGAGAGCAGCCAAAC 185
Db 1161 CACAGAACACTAATAATAACAGCAACGGGAACATATGTAGGTGGCAGCCACAAACAGAG 1220
QY 186 ACCACGATACCGGAAACAGACTCATGTTGACCTTGTAAAGAGAGAGAGCGGTTA 245
Db 1221 GCTAACTTAGTGAGAGAAATCATGAGAAATACGACAAAGATGAGATGAGGTGTT 1280
QY 246 TTATTTTC 253
Db 1281 TTCTTATC 1288

RESULT 4

US-09-793-807-10
; Sequence 10, Application US/09793807
; Publication No. US20030171310A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL EXPRESSION
; FILE REFERENCE: RFS-0206
; CURRENT APPLICATION NUMBER: US/09/793,807
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 35000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
US-09-793-807-10

Query Match 9.4%; Score 37; DB 10; Length 35000;
Best Local Similarity 49.2%; Pred. No. 6.8; Mismatches 100; Indels 0; Gaps 0;
Matches 97; Conservative 0;

QY 192 ATAACGGGAAACAGACTCATGTTGACCTTGGTGTAAAGAGAGAGAGCGGTTATTTT 251
Db 3747 ATTAGTGTCAAAAGAACCAATGTTGTTCTTATTGAAAGGTCACTGCAATATATAATTTAC 3806
QY 252 TCAGCATCTGTCGCCAGAGAGGCGCATGGAAGCCGGCGAGAGACATTTGCTGTAG 311
Db 3807 TTTCTAACATTTTCCGTGCTTAGGAATGACAGATTACTGTTGAGTGAACATCATTAATAG 3866
QY 312 ATTGATATTTAATATATTAGCGTAACCTGTTATGCTGTTATCTATATATTGTCATCTAAT 371
Db 3867 TACTATTTTAAAACTTGTGTCACCTGTTGCTTTTATTTTAGAAGAAATTTGGAT 3926
QY 372 CACTTTTAAGTCAGACT 388
Db 3927 TATTTTAGAACAAAT 3943

RESULT 5

US-10-311-455-2044
; Sequence 2044, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander

APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2044
LENGTH: 13503
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2044

Query Match 9.4%; Score 36.8; DB 15; Length 13503;
Best Local Similarity 63.6%; Pred. No. 5;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 299 ACATTGCTGTAGATGATTTATATATATAGCGTAACCTGTATGCTGTTATCTATATT 358
DB 10532 AATATTGATGAAGTAGATTTTATTTAATAGTGTAATTTTAGTTTATTTTAAATT 10591

QY 359 ATGTGATCTAAATCACTTTTAAGTCAGA 386
DB 10592 TAGTTAATAAATTTAGTGTAAATTCGA 10619

RESULT 6
US-09-864-761-6274
Sequence 6274, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aemica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SEQ ID NO 6274
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
LENGTH: 397
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004056.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
US-09-864-761-6274

Query Match 9.3%; Score 36.6; DB 9; Length 397;
Best Local Similarity 47.2%; Pred. No. 1.1;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 156 AAAAATATCGTGGGAAGGACGCCAAACACCCAGCATAACGGGAACACAGACTCATGTTG 215
DB 135 AAAAATATCGTGGGAAGGAAATGAAAGATCTGGCTCTTGGTCAACAGCAAAAGACA 194

QY 216 ACCTTGGTTGTAAGAGAGAGAGCGGCGCTTATTTTTCAGCATCTGTCGCGGAGAGAAG 275
DB 195 AACCTTCATCTAAAGTGATATAATAGTTATTACATACATCAATGAAGTCCAGCCTTGTATC 254

QY 276 GGCATGGAAGCCGGCGAGAGCAACATTGCTGTAGATTGATATTATTAATATTAGCGTA 335
DB 255 TGCTGATAGATGCTGTGTGTTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 314

QY 336 ACTGTTATGCTGTTATCTATATATTATGTGATCTAAATCACTTTTAAGTCAGAGTGA 390
DB 315 TAGAGAATGACACCTGGAGGAGATAATAATGTTAACACAATTTAAGTCATATAGA 369

RESULT 7
US-10-087-192-1498/c
Sequence 1498, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: CANCER
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1498
LENGTH: 330973
TYPE: DNA
ORGANISM: Homo sapiens

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(330973)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1498

Query Match          9.3%; Score 36.6; DB 13; Length 330973;
Best Local Similarity 47.2%; Pred. No. 26;
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 156 AAAAATATCGTGGGAGGACGACCCCAACACACGATACCGGAAACAGAGCTCATGTTG 215
Db 254607 AAAAATATGTTTAAAGAAATGAATGAAGATCTGGCTCTTGTCACAGGCAAGACA 254548

QY 216 ACCTTGTTGTAAGAGAGAGCAGGCGTTATTATTTTCAGCATCTGTGCGCCGACAGAG 275
Db 254547 AACCTTCATCAAAGTGATATAAATAGTTATTACATACATCATTAAGCTCCAGCCTTGATC 254488

QY 276 GGCATGGAAGCGCGGAGAGCAACATTGCTGTAGATTGATTTAATATATTAGCGTA 335
Db 254487 TGTGATAGAGTCTGTGTGGTCTTCATCCAGAGGGTGGATACCCCAAGTAGAATA 254428

QY 336 ACTGTTATGCTGTATCTATATTATATGATGATCATCTAAATCACTTTTAAGTCAGAGTGA 390
Db 254427 TAGAGATGACACCTGGAGGAGATATATGTTAACACAAATTTAAGTCATATAGA 254373

RESULT 8
US-10-335-977-4100
; Sequence 4100, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4100:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...714
; SEQUENCE DESCRIPTION: SEQ ID NO: 4100:
US-10-335-977-4100
```

```
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...711
; SEQUENCE DESCRIPTION: SEQ ID NO: 4100:
US-10-335-977-4100

Query Match          9.1%; Score 35.6; DB 16; Length 711;
Best Local Similarity 51.9%; Pred. No. 2.9;
Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 30 AAATAAAACAGCAGCAATAGAGCGGGTGTTCGAACAACCGCCGATGCCAGACAAAGTCG 89
Db 344 AATCATCTCCTTAGACCAAAAGACTGCTGAGCAACAATGCGGAATATCCGATAAGTTT 403

QY 90 TAGATACCTTTGGCAACAAGTACCGGCGCACCGACGAGAAAATCGCAGTGATAAAAGC 149
Db 404 TAGTTCTTTTATTGTGAGGGAGAGCTGAAAATCGAGCAAGAAAACAAAATATAAGAGAGC 463

QY 150 AATTCAAAAATATATCGTGGGAAAGGACAGCCAA 183
Db 464 GTTTAGAAACTTCTTAAGGCATATCAGAGCAA 497

RESULT 9
US-10-335-977-4099
; Sequence 4099, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4099:
SEQUENCE CHARACTERISTICS:
LENGTH: 714 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...714
; SEQUENCE DESCRIPTION: SEQ ID NO: 4099:
US-10-335-977-4100
```

US-10-335-977-4099

Query Match 9.1%; Score 35.6; DB 16; Length 714;
 Best Local Similarity 51.9%; Pred. No. 2.9; Mismatches 0; Gaps 0;
 Matches 80; Conservative 0; Indels 74; Indels 0; Gaps 0;
 QY 30 AATAAACAAGCAGCAATAGACGCGGTGTTGAACAACCGCGATGCCAGACAAAGTCG 89
 Db 344 AATCAATCCCTTAGACCAAAAGACTGCTGAGCAACAATCGGAATATCCGATAAAGTTT 403
 QY 90 TAGATACCTGTTGGCACAAGTACCGCGCCACAGCCAGCAAAATCCCGATGATAAAGC 149
 Db 404 TAGTTCTTTATTGTGAGGGAAGCTGAAATCGAGCAAGAAACAAAATATAAGAGAGC 463
 QY 150 AATCAAAAATATCGTGGGAGGACAGCCCAA 183
 Db 464 GTTAGAAACTTCTCTAAAGCATATCAGAGCAA 497

RESULT 10

US-10-335-977-4101
 ; Sequence 4101, Application US/10335977
 ; Publication No. US20040052799A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DOUGLAS SMITH et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; RELATING TO HELICOBACTER PYLORI FOR
 ; DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 10031
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: LAHIVE & COCKFIELD
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109-1875
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: UNIX
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/335,977
 ; FILING DATE: 30-Dec-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/993,002
 ; FILING DATE: 17-DEC-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mandragoras, Amy E.
 ; REGISTRATION NUMBER: 36,207
 ; REFERENCE/DOCKET NUMBER: GTN-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617)227-7400
 ; TELEFAX: (617)742-4214
 ; INFORMATION FOR SEQ ID NO: 4101:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 723 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Helicobacter pylori
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (B) LOCATION 1...723
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4101:
 US-10-335-977-4101

Query Match 9.1%; Score 35.6; DB 16; Length 723;
 Best Local Similarity 51.9%; Pred. No. 2.9;

Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 30 AATAAACAAGCAGCAATAGACGCGGTGTTGAACAACCGCGATGCCAGACAAAGTCG 89
 Db 356 AATCAATCCCTTAGACCAAAAGACTGCTGAGCAACAATCGGAATATCCGATAAAGTTT 415
 QY 90 TAGATACCTGTTGGCACAAGTACCGCGCCACAGCCAGCAAAATCCCGATGATAAAGC 149
 Db 416 TAGTTCTTTATTGTGAGGGAAGCTGAAATCGAGCAAGAAACAAAATATAAGAGAGC 475
 QY 150 AATCAAAAATATCGTGGGAGGACAGCCCAA 183
 Db 476 GTTAGAAACTTCTCTAAAGCATATCAGAGCAA 509

RESULT 11

US-09-754-853A-4/c
 ; Sequence 4, Application US/09754853A
 ; Publication No. US20030005491A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hauge, Brian M.
 ; APPLICANT: Parnell, Laurence D.
 ; APPLICANT: Parsons, Jeremy D.
 ; APPLICANT: Wang, Ming Li
 ; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
 ; FILE OF INVENTION: Soybean Cyst Nematode Resistance
 ; FILE REFERENCE: 38-10(15810)B
 ; CURRENT APPLICATION NUMBER: US/09/754,853A
 ; CURRENT FILING DATE: 2001-01-05
 ; PRIOR APPLICATION NUMBER: US 60/174,880
 ; PRIOR FILING DATE: 2000-01-07
 ; NUMBER OF SEQ ID NOS: 1119
 ; SEQ ID NO 4
 ; LENGTH: 513509
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (111805)..(113968), (114684)..(115204)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(513509)
 ; OTHER INFORMATION: unsure at all n locations
 ; OTHER INFORMATION: Clone ID: 318013_region_A3
 US-09-754-853A-4

Query Match 8.9%; Score 35; DB 10; Length 513509;
 Best Local Similarity 47.9%; Pred. No. 1e+02;
 Matches 101; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 182 AATCACCAGTAAACGGGAACAGACTCATGTTGACCTTGTTGTTAAAGAGAGAGGAGGC 241
 Db 275036 AATAACAAATCAACCCGAAATATAAACAATTAAGCTTAAGCTTAAATTCAAAAA 274977
 QY 242 GTTATTATTTTCAGCATCTCTCGCCGAGAGAAGGGCATGAAAGCGCGGAGAGCAAC 301
 Db 274976 ATATTATTATTATTTTATAAGCAAAAAAGGCAATGATATATTGATAAAATGAT 274917
 QY 302 ATTGCTGTAGATTGATATTAAATATATTAGCGTAACGTGTTATGCTGTTATCTATTATG 361
 Db 274916 ATATATGTATATAGATCAATAATTTTGCACAAAAATCTTAACATGTAACCTCTTTAT 274857
 QY 362 TGATCTAAATCACTTTTAAAGTCAGAGTGAAT 392
 Db 274856 TTCTTTAAATTCACITTAGGCTCTGTTGAT 274826

RESULT 12

US-10-433-793-124/c
 ; Sequence 124, Application US/10433793
 ; Publication No. US20040142334A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Epigenomics AG

```

; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 124
; LENGTH: 83391
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (79052..79053, 79055..79056, 79058..79059, 79061..79062, 79064..79065)
; OTHER INFORMATION: unknown base
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (79067..)
; OTHER INFORMATION: unknown base
; US-10-433-793-124

Query Match      8.8%; Score 34.4; DB 17; Length 83391;
Best Local Similarity 56.0%; Pred. No. 65;
Matches 65; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY    278 CATGAAGCCGGCGAGAGAACAATTCTGTGATGATGATATTTAATATATTAGCGTAAC 337
DB    24919 CATAAATCCACATATTTCAATACGTTTTATCTAAATATATATACTTCTCATAC 24860

QY    338 TGTATGCTGTTATCTATATTTATGTCATCTAAATCACCTTTTAAGTCAGAGTGAAATA 393
DB    24859 CATTATAAACCTTATTTCCATTATATTATATAAATAACTATACTATTATAAAAAA 24804

RESULT 13
US-10-027-632-262484/c
; Sequence 262484, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262484
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-262484

Query Match      8.7%; Score 34; DB 13; Length 1111;
Best Local Similarity 54.9%; Pred. No. 11;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY    112 CCCGGCGCACCAGCCAGAAAATGCCAGTGATAAAAGCAATTCAAAAATATCGGTGGGA 171

```

Db	193	CCTGGGCAACCAAGAGCGAACTCTGTCTCGAAAAA	AAAAAAAAAAAAAAAAAATCAGAGACA	133
QY	172	AGGACAGAGCCCAACACCAAGCAATACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGA	231	
Db	133	GGAGTAATAATAAACCTCAGTAGTGAAGGAAAAA	CCATGAATGCTGACATTGTTATTAAAAA	74
QY	232	GA	233	
Db	73	TA	72	

RESULT 14
 US-10-027-632-262484/c.
 ; Sequence 262484, Application US/10027632
 ; Publication NO. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827,129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 262484
 ; LENGTH: 1111
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-262484

Query Match	8.7%	Score 34;	DB 15;	Length 1111;
Best Local Similarity	54.9%	Pred. No. 11;		
Matches 67;	Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0

QY	112	CCCGCGGCACACGACCCAGAAATCCAGTGTAAAGCAATTCGGAATATCGGTGGCA	171	
Db	193	CCTGGGCAACCAAGAGCGAACTCTGTCTCGAAAAA	AAAAAAAAAATCAGAGACA	134
QY	172	AGGACAGAGCCCAACACCAAGCAATACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGA	231	
Db	133	GGAGTAATAATAAACCTCAGTAGTGAAGGAAAAA	CCATGAATGCTGACATTGTTATTAAAAA	74
QY	232	GA	233	
Db	73	TA	72	

RESULT 15
 US-10-289-762-1/c
 ; Sequence 1, Application US/10289762
 ; Publication NO. US20040006218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/10/289,762
 ; CURRENT FILING DATE: 2003-03-27

Db	193	CCTGGGCAACCAAGAGCGAACTCTGTCTCGAAAAA	AAAAAAAAAAAAAAAAAATCAGAGACA	133
QY	172	AGGACAGAGCCCAACACACGAGTAACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGA	231	
Db	133	GGAGTAATAATAAACCTCAGTAGTGAAGGAAAAA	CCATGAATGCTGACATTGTTATTAAAAA	74
QY	232	GA	233	
Db	73	TA	72	

RESULT 14
 US-10-027-632-262484/c.
 ; Sequence 262484, Application US/10027632
 ; Publication NO. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827,129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 262484
 ; LENGTH: 1111
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-262484

Query Match	8.7%	Score 34;	DB 15;	Length 1111;
Best Local Similarity	54.9%	Pred. No. 11;		
Matches 67;	Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0

QY	112	CCCGCGGCACAGCCAGCAATATCCAGTGTAAAGCAATTCGAAATATCGGTGGCA	171	
Db	193	CCTGGGCAACCAAGAGCGAACTCTGTCTCGAAAAA	AAAAAAAAAATCAGAGACA	134
QY	172	AGGACAGAGCCCAACACACGAGTAACGGGAAACAGACTCATGTTGACCTTGGTTGTAAGA	231	
Db	133	GGAGTAATAATAAACCTCAGTAGTGAAGGAAAAA	CCATGAATGCTGACATTGTTATTAAAAA	74
QY	232	GA	233	
Db	73	TA	72	

RESULT 15
 US-10-289-762-1/c
 ; Sequence 1, Application US/10289762
 ; Publication NO. US20040006218A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Griffais, R.
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides,
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis
 ; TITLE OF INVENTION: and treatment of infection
 ; FILE REFERENCE: 9710-003-999
 ; CURRENT APPLICATION NUMBER: US/10/289,762
 ; CURRENT FILING DATE: 2003-03-27


```

/ NUMBER OF SEQ ID NOS: 6849
/ SEQ ID NO 1
/ LENGTH: 1230025
/ TYPE: DNA
/ ORGANISM: Chlamydia pneumoniae
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(15000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (15001)..(30000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (30001)..(45000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (45001)..(60000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (60001)..(75000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (75001)..(90000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (90001)..(105000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (105001)..(120000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (120001)..(135000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (135001)..(150000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (150001)..(165000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (165001)..(180000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (180001)..(195000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (195001)..(210000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (210001)..(225000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (225001)..(240000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (240001)..(255000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (255001)..(270000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (270001)..(285000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (285001)..(300000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (300001)..(315000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (315001)..(330000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (330001)..(345000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (345001)..(360000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (360001)..(375000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (375001)..(390000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (390001)..(405000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (405001)..(420000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (420001)..(435000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (435001)..(450000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (450001)..(465000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (465001)..(480000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (480001)..(495000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (495001)..(510000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (510001)..(525000)
/ OTHER INFORMATION: n=a or c or g or t
/ FEATURE:

```

```
; NAME/KEY: misc feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; FEATURE:
Query Match      8.7%   Score 34; DB 16; Length 1230025;
Best Local Similarity 52.9%; Pred.No. 3e+02;
Matches 73; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 79 AGACAAAGTCGTAGATACCTGTTGGCCACAAGTACC GGCGGCCAC CAGCGCAAAAATCGCCA 138
Db 1176718 AGAGAAAAAATCTGTAGCAAATTGGCAAGAATAGCGGTTAACAGCAAGCAAAAATTCCA 1176659
QY 139 GTGATAAACAATTCAAAAAATATCGTGGAAGGACACCCCCCAACACACCACCATTAACGG 198
Db 1176658 ATTAAGAACGTATTCGAAGAGATATCGATTTTTAGTCMAAATATGAGGAATCTCGCAATGG 1176599
QY 199 GAACACGACTCATGTTGA 216
Db 1176598 TAACGATCCTCACCATGA 1176581
```

Search completed: December 18, 2004, 23:07:19
Job time : 279.694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 18, 2004, 22:46:01 ; Search time 1710.72 Seconds
(without alignments)
8371.204 Million cell updates/sec

Title: US-10-759-889-3

Perfect score: 393

Sequence: 1 ttcaacctcaaacgacagt.....ctttaagtcagagtgaata 393

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288.2	73.3	755	8	AF094919
2	162.8	41.4	785	9	CL663057
3	40.2	10.2	942	9	AL272552
4	39.6	10.1	1101	9	AL098287
5	39	9.9	790	6	CB665130
6	38.4	9.8	424	5	BU673561
7	38.4	9.8	453	6	CF324862
8	38.4	9.8	727	2	BB657476
9	38.4	9.8	846	9	CN8033M4
10	38	9.7	707	6	CD849459
11	38	9.7	755	8	BH485998
12	37.8	9.6	417	1	AL389174
13	37.8	9.6	524	9	CE225599
14	37.8	9.6	3190	3	AK085573
15	37.2	9.5	603	8	AZ230411
16	37.2	9.5	665	4	BG546513
17	37.2	9.5	997	9	CN8005TE
18	36.8	9.4	922	9	CN80073W
19	36.6	9.3	296	7	CO253598
20	36.6	9.3	812	9	CC552656
21	36.6	9.3	822	8	BZ603536
22	36.4	9.3	616	9	CE401841
23	36.4	9.3	861	8	AZ528603
24	36.2	9.2	406	1	AL389175

25	36.2	9.2	593	5	BO702511
26	36.2	9.2	840	9	AG532072
27	36.2	9.2	1000	7	CR278250
28	36	9.2	652	9	CB662916
29	35.8	9.1	928	9	CNS0020M
30	35.8	9.1	293	2	BB193716
31	35.8	9.1	665	9	CE011345
32	35.6	9.1	238	6	CB279545
33	35.6	9.1	675	9	CC946844
34	35.6	9.1	677	8	BZ089375
35	35.6	9.1	756	8	BZ451202
36	35.6	9.1	1139	8	CC220232
37	35.6	9.1	1302	8	BZ556775
38	35.2	9.0	679	5	BU484568
39	35.2	9.0	731	8	BZ049478
40	35.2	9.0	867	7	CR289354
41	35.2	9.0	1044	9	CL503018
42	35	8.9	895	5	BU152641
43	35	8.9	913	7	CO081762
44	35	8.9	922	7	CK016766
45	35	8.9	1003	4	BG179376

ALIGNMENTS

RESULT 1
AF094919
LOCUS AF094919 755 bp DNA linear GSS 29-AUG-2000
DEFINITION AF094919 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 906-T3, genomic survey sequence.
ACCESSION AF094919
VERSION AF094919.1 GI:4322761
KEYWORDS GSS.
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 755)
AUTHORS Wong, R.M.Y. and McClelland, M.
TITLE End Sequences of Salmonella typhimurium LT2 Lambda DASHII Clones, Li-Cor
JOURNAL Unpublished (1999)
COMMENT Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.

FEATURES

source
Location/Qualifiers
1..755
/Organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone="906-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor sequencer"

ORIGIN

Query Match 73.3%; Score 288.2; DB 8; Length 755;
Best Local Similarity 85.0%; Pred. No. 1.8e-75;
Matches 334; Conservative 0; Mismatches 58; Indels 1; Gaps 1;
QY 1 TTCAACTCAACGAACAGTCGGGATATCAATAAACAAGCAGCAATAGAGCGGTGT 60
DB 189 TTCAACTCAACGAACAGTCGGGATATTAATAAACAAGCAGCAATAGAGCGGTAT 248
QY 61 TGAACAACGGCGGATGCCAGCAAAAGTCGTAGATACCTTTGGCACAAGTACCGGGGCA 120
DB 249 TGAACAGAGCGGATGCCAGCAAAATCATAGATGCCCGTCGCAACGATCCGACGCA 308

[illegible]

JOURNAL
COMMENT

Unpublished (2002)
Contact: Reddy AR
Department of Plant Sciences, School of Life Sciences
University of Hyderabad
P.O. Central University, Hyderabad-500 046, A.P., India
Tel: 0091-40-3010265
Fax: 0091-40-3010145
Email: arjuls@uohyd.ernet.in
Insert Length: 424 Std Error: 0.00
Seq primer: CGCAGGCTTTCCAGTCACGAC.

FEATURES

source

1. 424
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nagina 22"
/db_xref="taxon:39946"

/clone="NL 17 12"
/tissue_type="Entire leaf tissue"
/dev_stage="35 day-old seedlings"
/clone_lib="Drought stress (leaf)"
/note="Organ: Leaf; Vector: T713pac; ESTs from normalized leaf cDNA Library from drought stressed seedlings"

ORIGIN

Query Match 9.8%; Score 38.4; DB 5; Length 424;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 72; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 126 CAGAAATGCCAGTGTATAAAGCAATTCAAAATATCGGTGGGAGGACAGCCCAAC 185
|||
DB 2 CACAAGAACACTAATAACAGCGGGAACATGTAGTGGCGGACCAACACGAG 61
|||
QY 186 ACCACGATACGGGAACAGACTCATGTGACCTTGGTTGTAAGAGAGAGAGCGGTGA 245
|||
DB 62 GCTAATCTTAGTGAGAGAGATCATGAGAGAGTACGACAAAGATGATGAGGTGTC 121
|||
QY 246 TTATTTTC 253
|||
DB 122 TTCTTATC 129
|||

RESULT 7

CF324862/c

LOCUS

DEFINITION CF324862 453 bp mRNA linear EST 18-AUG-2003
JMT1--01-K19.g1 AtJMT-overexpressing transgenic rice lambda phage
cDNA library (JMT1) Oryza sativa (japonica cultivar-group) cDNA

clone JMT1--01-K19, mRNA sequence.

ACCESSION

VERSION

CF324862.1

GI:33798004

EST.

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoidae; Oryzaceae; Oryza.

1 (bases 1 to 453)

Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,

Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongul University

Yongin, Kyonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

Location/Qualifiers

1. 453

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nackdong"

/db_xref="taxon:39947"

/clone="JMT1--01-K19"

/tissue_type="leaf"

/dev_stage="14 days after germination"

/lab_host="E.coli SOLR"

/clone_lib="AtJMT-overexpressing transgenic rice lambda

phage cDNA library (JMT1)"

/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:

XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5'

XhoI and 3' end with EcoRI and 3' end with XhoI site. mRNA was

prepared from Arabidopsis Jasmonate Carboxyl

methyltransferase overexpression line."

ORIGIN

Query Match 9.8%; Score 38.4; DB 6; Length 453;
Best Local Similarity 56.2%; Pred. No. 2.4;
Matches 72; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 126 CAGAAATGCCAGTGTATAAAGCAATTCAAAATATCGGTGGGAGGACAGCCCAAC 185
|||
DB 448 CACAAGAACACTAATAACAGCGGGAACATGTAGTGGCGGACCAACACGAG 389
|||
QY 186 ACCACGATACGGGAACAGACTCATGTGACCTTGGTTGTAAGAGAGAGAGCGGTGA 245
|||
DB 388 GCTAATCTTAGTGAGAGAGATCATGAGAGAGTACGACAAAGATGATGAGGTGTC 329
|||
QY 246 TTATTTTC 253
|||
DB 328 TTCTTATC 321
|||

RESULT 8

BB657476/c

LOCUS

DEFINITION BB657476 727 bp mRNA linear EST 26-OCT-2001
musculus cDNA clone D230002J12 5', mRNA sequence.

ACCESSION

VERSION

BB657476.1

GI:16491302

EST.

KEYWORDS

SOURCE

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 727)

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,

Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,

Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,

Ono,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,

Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho,

Tsukuba, Ibaraki, 305-8565, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,

Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.

and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,

Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
source
1. .727
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230002J12"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo
eyeball"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGCGCGCAACTCGAGTGTGTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATCTCGATTAAATTAATATCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBlueScript KS(+) after bulk excision from Lambda FLC I."

ORIGIN
Query Match 9.8%; Score 38.4; DB 2; Length 727;
Best Local Similarity 49.5%; Pred. No. 2.6;
Matches 99; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 32 ATAAACAGAGCGAATAGAGCGCGGTGTGTAACAAACCGCGGATCCAGACAAAGTCGTA 91
Db 522 ATACAACTAACAAAATAAGATAGAGTAAATAAAACAGAGTAGAGACGATATATTA 463
QY 92'GATACCTGTGGCACAAGTACCGCGGCACCGCCAGAAATCGCGAGTGTAAAGCAA 151
Db 462 GAAATAAATATAGATAGTAGGAGGGGAGAGAGGAATATAAATATAAATAATTGAA 403
QY 152 TTCAAAAAATATCGGTGGGAAGGACAGCCCAACACACGATACCGGGAACAGACTCAT 211
Db 402 TTCAATAAGTACAGCTTGAATAATAAACAGAGTTTAAAAAAAACCTGTAGCTCACCTAT 343
QY 212 GTTGACCTTGGTTGTAAGA 231
Db 342 AATCACCTCGTTTATAAACA 323

RESULT 9
CNS033M4 846 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION 209C18 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL226309
VERSION AL226309.1 GI:7885226
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoideae; Tetraodontidae; Tetraodon.

REFERENCE

1
AUTHORS
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fzames, C., Wincker, P., Brotier, P., Quetier, P.,
Saurin, W. and Weissenbach, J.

TITLE
Estimate of human gene number provided by genome-wide analysis

JOURNAL
using Tetraodon nigroviridis DNA sequence

MEDLINE
20296633

PUBMED
10835645

REFERENCE

2
AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, P.,
Saurin, W., Bernot, A. and Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the

JOURNAL
freshwater pufferfish Tetraodon nigroviridis

MEDLINE
20359837

PUBMED
10899143

REFERENCE
3 (bases 1 to 846)

Genoscope.

Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

<http://www.genoscope.cns.fr/Tetraodon>.

Location/Qualifiers

1. .846

FEATURES

source

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone_lib="G"

/note="Genoscope sequence ID : COAG209BB09SP1-end ;

PUC-Ori"

ORIGIN

Query Match 9.8%; Score 38.4; DB 9; Length 846;
Best Local Similarity 40.9%; Pred. No. 2.7;
Matches 81; Conservative 23; Mismatches 94; Indels 0; Gaps 0;
QY 9 CAAACGAAACAGTCGCGATATCAATATAAACAAGCAGCAATAGAGCGCGGTGTTGAACAAC 68
Db 213 CAAARAAAAAAGAAAGAAAGACGAGARGGMAAAACCCARAGWAGWAGAAARAG 272
QY 69 GCGGATGCCAGACAAAGTCGTAGATACCTGTTGGCACAAGTACCGCGCGCACCGCCAG 128
Db 273 GAGGAGCAGACAMWAGAGMGCGCAGMARGMAGGGGCGMACAMMGCAACGCGCAGGAGAA 332
QY 129 AAAATCCGAGTCGATAAAAGCAATTCAAAAAATATCGTGGGAGGACAGCCCAACACC 188
Db 333 AAAMAGCMCAGAAACCCAGGCGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 392
QY 189 ACCGATACGGGAAACAGCA 206
Db 393 AAGAAAACRAGCARAACA 410

RESULT 10
CD849459/c 707 bp mRNA linear EST 11-JUL-2003
LOCUS DRCAC020D12FM1 HadevR2 Helianthus annuus cDNA clone HadevR2020D12,
DEFINITION mRNA sequence.
ACCESSION CD849459
VERSION CD849459.1 GI:32533281
KEYWORDS EST.
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.

REFERENCE

1 (bases 1 to 707)

AUTHORS

Genoplante.

JOURNAL

Unpublished (2003)

COMMENT

Contact: Genoplante

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french

plant genomics programme 'genoplante' (<http://www.genoplante.com>

and <http://genoplante-info.infobiogen.fr>).

FEATURES

source

1..707

/organism="Helianthus annuus"

/mol_type="mRNA"

/cultivar="psc8"

/db_xref="taxon:4232"

/clone="HadeVR2020D12"

/tissue_type="terminal bud"

/clone_lib="HadeVR2"

ORIGIN

Query Match

Best Local Similarity 9.7%; Score 38; DB 6; Length 707;

Matches 92; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 112 CCGCGCGCCACCGCAGAAATGCGAGTGAATAAGCAATTCAAAAATATCGGTGGGA 171

DB 564 CCGAAGTCAAGGAAACAAAGACGGCGGTTCACATAGCAAACTACATCGGAGCCA 505

QY 172 AGACAGAGCCCAACACACAGTAAAGGAAAGACGACTCATGTGACCTGGTGTAAAGA 231

DB 504 TGGGCAACCCACGTAACACACAGACAGAAAGATAGACATGCGGACCGGTGATACGA 445

QY 232 GAGAGCAGGCGTTATTATTTCAGCATCTGTCGCGCAGAGAGGCGCATGGAAGCCGGG 291

DB 444 AGSCGAGAGATAGACATGACGCGCGCGTGTACTAAGAGAGCGGAGACGGGTGA 385

QY 292 CGA 294

DB 384 CGA 382

RESULT 11

BH485998/c

LOCUS

BOHIH37F BOHI Brassica oleracea genomic clone BOHIH37, genomic

survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 755)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished (2001)

Other_GSSs: BOHIH37R

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..755

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOHIH37"

/clone_lib="BOHI"

/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN

Query Match

Best Local Similarity 9.7%; Score 38; DB 8; Length 755;

Matches 56; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 301 CATTCCTGTAGATGATTAATATATATAGCTAACTGTTATGCTGTATCTATATTAT 360

DB 234 CATACACGTATATTAATTAATTTAATAATAATAAGTAAATATTTATTTATTTAT 175

QY 361 GTGATCTAAATCACTTTTAAGTCAGA 386

DB 174 AGGTTTAAATCTTTTATTTATA 149

RESULT 12

AL389174

LOCUS

MCBC53C03F1 MtBC Medicago truncatula cDNA clone MtBC53C03 T3, mRNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula (barrel medic)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

1 (bases 1 to 417)

Journet, B.P., Crespeau, H., van Tuinen, D., Gouzy, J., Jaillon, O.,

Niebel, A., Carreau, V., Chatagnier, O., Kahn, D.,

Gianinazzi-Pearson, V. and Gamas, P.,

Medicago truncatula ESTs from endomycorrhizal roots

Unpublished (2000)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqrf@genoscope.cns.fr Web : www.genoscope.cns.fr

Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de

Biologie Moleculaire des Relations Plantes-Microorganismes,

CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :

Mt-est@toulouse.inra.fr Website :

<http://sequence.toulouse.inra.fr/Mtruncatula.html>).

Location/Qualifiers

1..417

/organism="Medicago truncatula"

/mol_type="mRNA"

/cultivar="Jemalong"

/db_xref="taxon:3880"

/clone="MtBC53C03"

/tissue_type="arbuscular mycorrhiza"

/dev_stage="harvested 3 weeks post inoculation with Glomus

intraradices"

/clone_lib="MtBC"

/note="Vector: pBluescript pSK; Site 1: EcoRI; Site 2:

XhoI; M. truncatula sterilised seeds were germinated for

72h at 25 C, before transplanting into a 1/3 Epsilones soil

: 2/3 calcined Terragreen mix in the presence of oncon

root fragments colonized by the arbuscular mycorrhizal

fungus Glomus intraradices (Schenck & Smith, isolate

LPAB). The plants were watered every day and twice a week with a modified nutrient Long Ashton solution without phosphate but with a high level of nitrate. After 3 weeks RNA was extracted from whole root systems. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into Uni-zap XR vector from Stratagene and packaged using Gigapack Gold packaging extracts. Plasmids containing cDNA inserts were mass-excised from phage stocks using ExAseit helper phage and propagated in SOLR cells. Clone ordering and sequencing was performed by the Centre National de Sequençage (Genoscope, Evry, France). Note : EST may be of fungal origin."

ORIGIN

Query Match 9.6%; Score 37.8; DB 1; Length 417;
Best Local Similarity 48.0%; Pred. No. 3.6; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 117;

QY 156 AAAATATCGTGGGAGACAGCCCAACACGACGATACCGGAAACAGACTCATGTTG 215
DB 162 AAAATATCGGAGGTAAAGTACAGGCAAACTACAATAACAGGAAACACTCTCATGTTT 221
QY 216 ACCTCGTGGTAAAGAGAGAGAGCGGTTATATTTCAGCATCTGTGCGCCAGAGAG 275
DB 222 ACCCTACCGCTGGCACCATTACTCTTTGTGATGATCAACACATTGACGGATTGAA 281
QY 276 GCATGAAAGCGGCGGAGAGCAACATTCGTGTAGATTGATTTATATATTAGCGTA 335
DB 282 GACGATGGAGTAGTTACATTAATCAATAATTTTGTAGTTGAGATTTCTTTTCAATA 341
QY 336 ACTGTTATCGTGTATCTATATATTGATGATCTAAATCACTTTAA 380
DB 342 TAGTTATTTTAAATTTAATAATTATCAATTAATAAACAATTGAA 386

RESULT 13
CE225599/c

LOCUS
DEFINITION
tigr-gss-dog-17000338761286 Dog Library Canis familiaris genomic,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.

The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED

COMMENT

Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES

source

Location/Qualifiers
1..524
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood."

ORIGIN

Query Match 9.6%; Score 37.8; DB 9; Length 524;
Best Local Similarity 52.9%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 72;

QY 8 TCAACCAAGACAGTCGCGATATCAAAATAAAACAGCAGCAATAGAGCGCGTGTGAACAA 67
DB 153 TTACAGGACCTGAAACTACATGAAACAACAACTGATGAACCTGAAAGAGAAATAGACAA 94
QY 68 CGCCGGATGCCAGACAAGTCGTAGATACCTGTTGGCACAAGTACCCGGCGCACCACCA 127
DB 93 TTCATAATTTAGTCAGAGACTTATATTCCTCTCTCCATAGTTAAACAGGGTAAAGTAGACA 34
QY 128 GAAATCGCCAGTCGATATAAAGCAATTCAAAAA 160
DB 33 GAAATCAGTACCACTACAGAAACACATCAA 1

RESULT 14
AK085573/c

LOCUS
DEFINITION

Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D630042J08 product:unclassified, full insert
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
10349636

REFERENCE

AUTHORS
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE

Hayashida, K., Hayatsu, N., Hiranoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tgawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://phantom.gsc.riken.jp/

FEATURES

Location/Qualifiers

1..3190

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:D630042J08"

/db_xref="taxon:10090"

/clone="D630042J08"

/issue_type="kidney"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

1..3190

/note="unclassifiable"

misc_feature

1..3190

ORIGIN

Query Match 9.6%; Score 37.8; DB 3; Length 3190;

Best Local Similarity 47.8%; Pred. No. 5.4; Mismatches 0; Gaps 0;

Matches 111; Conservative 0; Indels 122; Indels 0; Gaps 0;

148 GCAATTCAAAATAATATCGGTGGGAAGCAGACGACGCCAAACACCAACGTAACGGAACACAC 207

1709 GGACTGAAACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1650

208 TCATGTTGACCTGTTGTTGAAGAGAGAGACGCGGTATTTATTTTCAGCATCTGCGCGG 267

1649 TGAAGAGGACATTCGTAGACACAGAAAGTAAACCTGTATTTATTTCTCGTGTGG 1590

268 CACAGAGGCGCATGGAAGACCGCGGAGAGCAACATTCGTGTAGATTGATATTTAATATA 327

1589 GAGTAGGACATGGAGTACACACACAGTTTACATAGTATGATGATGATGATGATGATGATG 1530

328 TTAGCGTAATGTTATGCTGTTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 380

1529 TCACATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1477

RESULT 15

AZ230411/c

LOCUS

DEFINITION

RPIC-23-69J14.TJ RPI-23 Mus musculus genomic clone RPI-23-69J14,

genomic survey sequence.

ACCESSION

AZ230411

VERSION

AZ230411.1

KEYWORDS

GSS.

SOURCE

Mus musculus

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mammalia, Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 603)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPI-23
Unpublished (1999)
Other GSSs: RPI-23-69J14.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPI-23. For BAC
library availability, please contact Pieter de Jong
pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Reseach Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plates: 69 rows; J column: 14
Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers

1..603

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPI-23-69J14"

/sex="female"

/lab_host="DH10B"

/clone_lib="RPI-23"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 9.5%; Score 37.2; DB 8; Length 603;

Best Local Similarity 53.4%; Pred. No. 5.8;

Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

210 ATGTTGACCTTGGTTGTAAAGAGAGAGAGCGGCGTTATTTTTCAGCATCTGTCGCCGA 269

273 ATGTTGAAATGGATTGTAGTCATATTGCAATGTTTTTCCCTGCAGCATCTAGTGCATC 214

270 GAGAAGGCGATGAAAGCGCGGCGAGAGCAACATTGCTGTAGATTGATTTAATATTT 329

213 AAAGTGCAGGAGGAAAGGAGGAGGACGTGCCAGAAATTAACCAATCTCTGTAATAATA 154

330 AGCGTAATCTGTTATCTGTTATCTAT 355

153 TGTGGAAGGTTAAGTAATTAAGTTAT 128

Search completed: December 19, 2004, 02:47:03

Job time : 1716.72 secs